Photo Forensics from Partial Constraints

A Thesis

Submitted to the Faculty

in partial fulfillment of the requirements for the

degree of

Doctor of Philosophy

in

Computer Science

by

Eric Kee

DARTMOUTH COLLEGE

Hanover, New Hampshire

April, 2013

Examining Committee:

(chair) Hany Farid, Ph.D.

Devin Balkcom, Ph.D.

Michael Black, Ph.D.

Lorenzo Torresani, Ph.D.

F. Jon Kull, Ph.D.
Dean of Graduate Studies

Abstract

A variety of forensic methods have been developed to identify falsified photos, each unified by the ability to estimate and detect properties of a photo that are perturbed by forgery. There exist, however, many photos in which the required properties cannot be estimated. We present an approach to detect forgery in these photos. We use this approach to detect physically inconsistent shadows and shading in photos for which it is not possible to estimate the associated lighting properties. Specifically, we develop a method to detect inconsistent shadows cast by point and area light sources when a strict shadow-to-object correspondence cannot be estimated. We further develop a method to detect inconsistencies between shadows and the shading on objects when object geometry is only partially known, and when objects are photographed under unknown perspective. We conclude by describing prior methods that can be generalized to analyze photos in which estimation is not possible.

Acknowledgments

I would like to first thank my advisor, Hany Farid. Hany brings tremendous and exemplary dedication and honesty to everything he does, and this extends to his students. I could not have written this dissertation without his guidance, patience, and personal investment in my training as a computer scientist. Hany was able to see potential in me as a young Ph.D. student, and when I did not see it myself. It is difficult to overstate his influence in my career and life.

I have also been fortunate to collaborate with professor James O'Brien on key portions of this dissertation. I would like to thank him for his many ideas, suggestions, and general banter. Our collaborations were both fruitful and fun.

I would further like to thank my committee members, Devin Balkcom, Michael Black, and Lorenzo Torresani for their helpful comments and suggestions during my proposal, their help in reviewing this dissertation, and their positive feedback. Thank you.

Beyond my committee, I would like to acknowledge the professors here at Dartmouth with whom I worked and studied. I was fortunate in my teaching assistantships to work closely with both Andrew Campbell and Sean Smith who later became mentors to me. I very much enjoyed working with them, and I would like thank them for their guidance and encouragement. I also greatly enjoyed the courses I took with Amit Chakrabarti, Hany Farid, Lisa Fleischer, Prasad Jayanti, Sean Smith, and Lorenzo Torresani. Each are talented and engaging lecturers.

I would like to offer a special thanks to the staff and system administrators, who consistently went out of their way to help me. I owe much of the computational results in this dissertation to the apparently tireless and enthusiastic support of Tim Tregubov, who made himself available to resolve problems with the departmental cluster at essentially any hour of any day. To Wayne Cripps, I owe thanks for his extensive lab configuration during my teaching assistantships. And, I thank Joseph Elsener for never once telling me that I should already know how to complete an expense report.

I owe many personal thanks to my family, who gracefully supported me from afar when my challenges were perhaps difficult to understand. I especially thank Stephanie, who supported me tirelessly and without question. I could not have reached this point without her.

Finally, I would like to close by thanking Jessica Fridrich, who once emphatically offered to me kind words that I feel every doctoral student should hear — *never give up.*

Contents

Chapter 1

Introduction

Photographic forgery has a history nearly as old as photography itself, with forgeries dating as early as the mid 1800s. Ironically, with the invention of the photograph to chronicle our visual history, photographic forgery soon provided an equally powerful tool for deception. The deceptive realism of such forgery was exploited by Stalin, Mao, Hilter, Mussolni, and Castro in attempts to alter history, and this deception continues today, Figure 1.1. Forgeries now influence nearly every aspect of our world from politics, to science, to law, and even to our self-image as impossibly perfect models routinely grace magazine covers. Yet modern forgeries are much easier to concoct than those in history. Digital technology has quickly placed into the hands of all, low-cost tools that have transformed the once expensive and challenging art of photo forgery. Tools that distinguish forgeries, however, lag behind. Such photo authentication tools are critical because they restrict the number of people capable of forgery, and can begin to restore our trust in photography.

1.1 Forensics

In the absence of an embedded watermark or digital signature with which to authenticate a photo, forensic methods assume that forgery will perturb properties of a photo that result from its digital processing, or from the physical characteristics of the optics and scene. Forgery can be detected by estimating these digital or physical properties, and measuring the extent to which they differ from their appropriate values.

In this section, we survey forensics from the perspective of the digital and physical properties that are estimated by forensic methods. An italicized summary is provided with each property for readers who wish to quickly survey the field. This summary is followed by an explanation of the methods that have been developed. For brevity, only the broader strategies, strengths, and weaknesses are described.

Prior surveys are roughly organized by property with only slightly less (and different) emphasis [199, 213, 170, 65, 12, 165, 137, 220, 180]. By emphasizing the estimated properties, we identify common challenges and strategies. References are listed in chronological order.

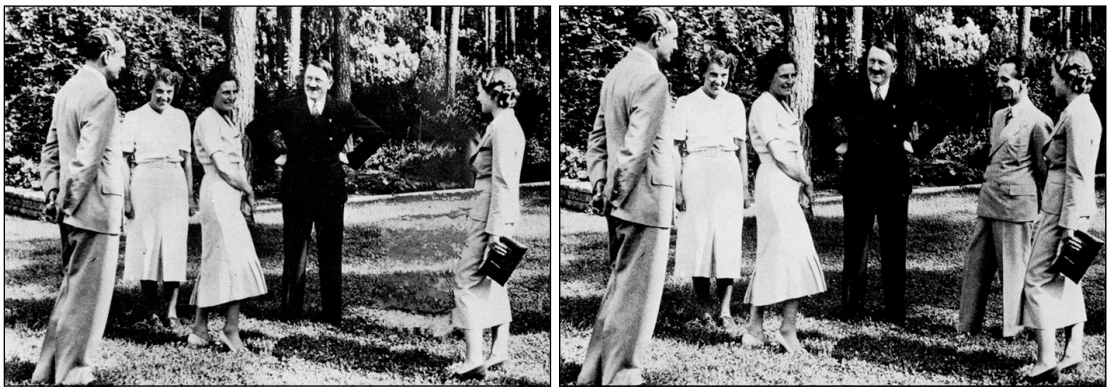


Figure 1.1: In this doctored photograph (left), taken in 1939, Adolf Hitler had Joseph Goebbels removed from the original photograph (right, second from the right). It remains unclear why exactly Goebbels fell out of favor with Hitler.

1.1.1 Digital properties

Modern cameras represent images digitally: light is digitized by an optical sensor, computationally transformed, and stored in a digital format. Subsequent photo modifications, which may constitute forgery, also operate within this digital domain. These digital processes give photos properties that are often visually subtle, and may be inadvertently modified, or created, by forgery. A diversity of forensic methods have therefore been proposed to estimate digital properties, and measure differences between those properties and their typical values.

Noise

Photos contain both scene content and various types of noise. In typical settings, the noise is visually subtle, both because it is suppressed by cameras, and because it is often disregarded by observers. When photos are modified, noise is nonetheless affected. For example, if a region of an image is brightened, noise may be amplified. Forgeries might therefore be detected by estimating noise properties, and measuring the extent to which they differ from their typical values.

A useful property of noise results from natural variations in the photo-sensitivity of individual pixels. This non-uniformity in pixel sensitivity imposes on the image a multiplicative noise pattern that is unique to every camera, and may be disturbed by forgery. For example, this noise pattern may be disturbed if a region is spliced into a photo. Photo tampering can be detected by estimating the noise pattern in a photo, and measuring the extent to which it differs from the typical pattern that should be observed. The typical noise pattern is unique to each camera, and must be pre-computed from the camera that produced the photo [76, 139, 43, 42, 262, 75, 34, 35]. Performance is degraded for cameras that perform lens correction, digital zoom, and compression, which distort the pattern; methods have

been developed to mitigate these issues [84, 83, 155]. A talented forger might insert fake noise [23, 214, 82], although some defenses exist against such attacks [86, 85].

When the camera-specific noise pattern of a photo is unknown, the level of noise can be assumed to be uniform across a photo. This uniformity may be violated when photos are falsified. For example, non-uniform levels of noise may be introduced if multiple photos are spliced together, or regions of a photo are brightened. Methods have therefore been proposed to estimate the noise level within regions, and to detect regions in which the noise level differs [195, 194, 171, 166, 89, 202]. Similarly, such generic noise properties have also been used to classify photos as computer generated or real [120, 53], and to detect brightness or contrast changes [58, 59].

Methods that measure differences in noise properties are attractive because techniques to estimate noise are well established. The use of photo-sensitivity noise is, however, only applicable when the camera is known, and inconsistent levels of noise may be visible to human observers. Further, methods that estimate noise properties are typically sensitive to compression and resizing of photos [196, 83].

Pixel correlation

Photos are encoded by discrete pixel values that may become correlated with their neighbors when photos are digitally processed. Such correlations are common but typically imperceptible. When photos are modified, however, pixel correlations may be destroyed or introduced. For example, if a photo is enlarged, pixel values must be interpolated, and this may introduce linear correlations between neighboring pixels. Forgeries might therefore be detected by estimating the extent to which neighboring pixels are correlated, and measuring when this correlation property differs from a typical photo.

One source of pixel correlation results from the process that cameras use to construct color images. Cameras sample color at alternating pixels in each color channel of an image, and must interpolate to recover the missing samples. This interpolation creates color samples that are correlated with their neighbors. Furthermore, the locations of these correlated pixels are periodic (because color samples alternate). When a photo is modified, these periodic correlations may be destroyed. For example, non-linear operations such as median filtering may destroy these correlations. Forgeries can therefore be detected by estimating the degree to which neighboring pixel values are related (correlated) according to known color interpolation algorithms. When these correlations are not periodic, photo modification is detected [71, 99, 55, 60, 100, 54, 205, 202, 121]. Similarly, it is possible to estimate the color interpolation coefficients, rather than the induced periodicity, as the digital property that evidences forgery [30, 243, 242, 242, 241, 239, 240]. The absence of color interpolation has been used to detect computer rendered images [55, 81].

Pixel correlations are also created by forgery, which often requires that regions of a

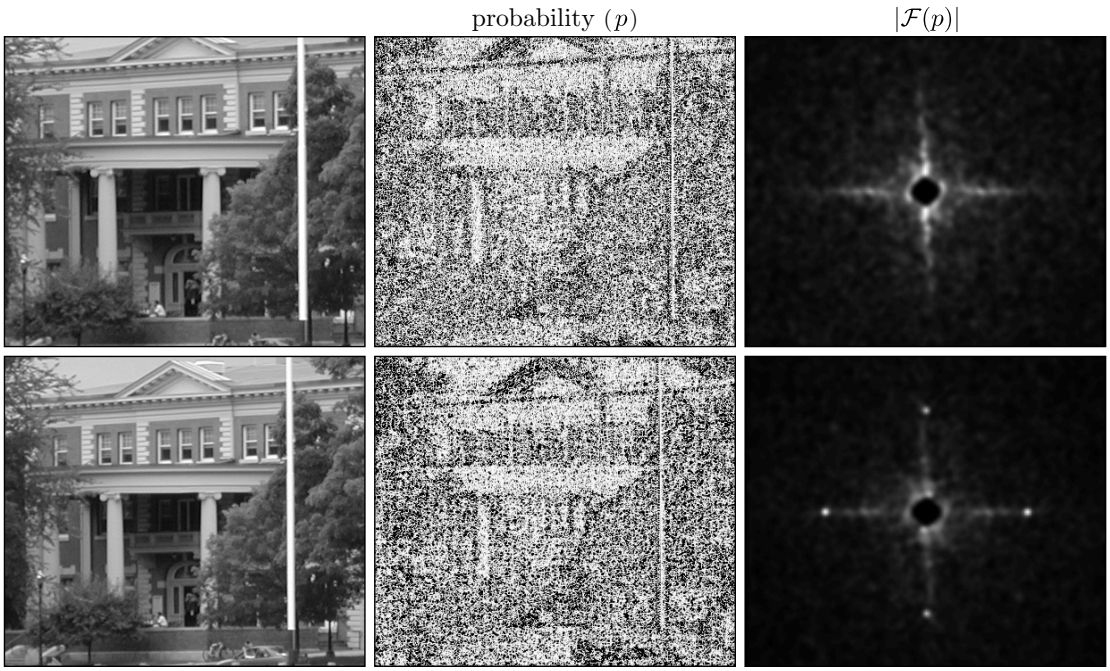


Figure 1.2: Pixel correlations are introduced when photos are resized, rotated, or otherwise resampled. Shown in the top row is an original photo (left), an estimation of the probability that each pixel is linearly correlated with its neighbors (middle), and the Fourier transform of this probability map (right). Shown in the bottom row is a photo that has been reduced by 10%. Resizing introduces periodic correlations between pixels, which introduces periodic patterns into the probability map (center). These periodic patterns can be detected as spikes in the Fourier domain (right). Figure due to Popescu and Farid [204].

photo be resized, rotated, or otherwise resampled. Resampling operations introduce periodic correlations that are distinct from those produced by camera color interpolation. Photo modifications that entail resampling can therefore be detected by estimating the periodicity of pixel correlations that result from resampling algorithms, and measuring when this periodicity is present in a photo [69, 26, 68, 146, 125, 130, 152, 169, 167, 123, 164, 200, 207, 80, 202, 206]. Figure 1.2 illustrates one such approach. Similarly, if resizing or rotation has been applied, it is possible to estimate the rotation or resizing values, rather than the induced periodicity of correlations, as the digital property that evidences photo modification [198, 51, 50, 52, 255, 254].

Pixel correlations are also created when photos are subjected to median filtering. Such filtering may be applied within a camera to suppress noise, or during forgery to hide photo manipulation. It is nonetheless useful in forensics to determine when median filtering has been performed. Median filtering increases the probability that adjacent pixels have the same value [19], which can be viewed as a correlation. Unlike resampling, correlations that result from median filtering are not periodic. Statistical properties of these correlations can however be estimated to detect median filtering [31, 32, 260, 29, 129]. Similarly, it is possible to estimate a linear approximation to the median filter as the digital property that

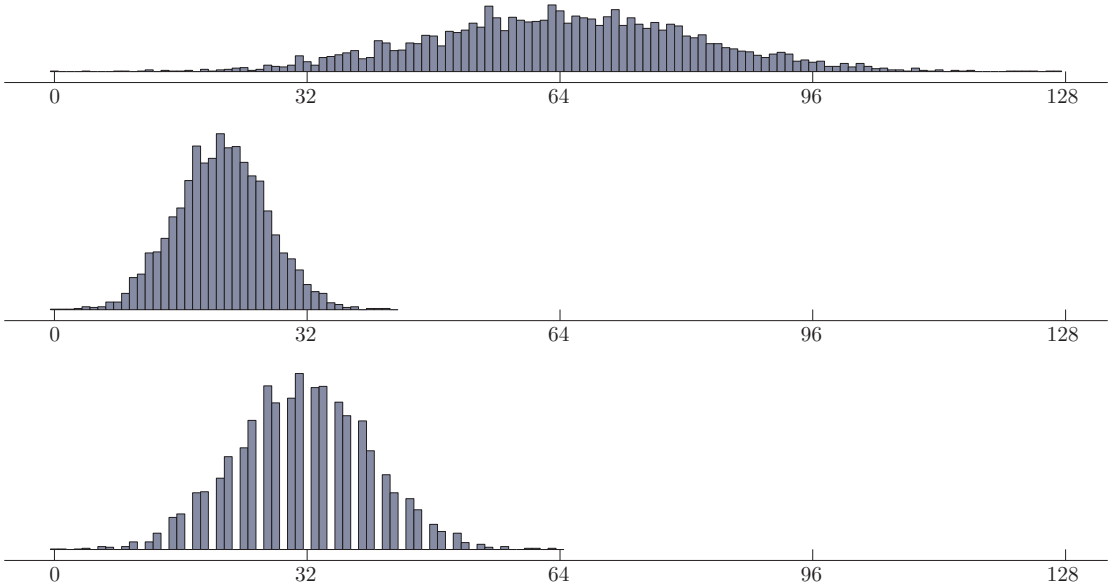


Figure 1.3: Distinct image statistics are introduced when a falsified JPEG photo is resaved in JPEG format. This double-JPEG compression introduces periodic artifacts into the histogram of DCT coefficients. Shown in the top row is a normally-distributed signal, which represents a hypothetical histogram of DCT coefficients. In a first JPEG compression, these coefficients are quantized by a factor of 3 (middle). In a second JPEG compression, the singly-quantized signal is expanded back to its original range of 128 and subsequently requantized by a factor of 2 (bottom). Because the second quantization factor is smaller than the first, periodic gaps are introduced; this periodicity can be measured as evidence of double-JPEG compression. Figure due to Popescu and Farid [202].

is used to detect forgery [243].

Methods based upon pixel correlation are attractive because this property is typically visually imperceptible, and pixel correlations are common in photos. Pixel correlations can be destroyed, however, by benign modifications such as JPEG compression or resizing [141, 124, 255, 177, 127, 82, 126, 164]. This difficulty has motivated more robust methods for detecting resampling [26, 130, 177]. In the case of color interpolation, it is desirable that inter-pixel correlations are easily destroyed, as the absence of such correlations evidences forgery. Color interpolation correlations can, however, be inserted during forgery to simulate an authentic photo [128].

Image statistics

Photos are not random collections of pixels, and are therefore governed by a variety of statistical properties. Such statistics may be difficult to visually identify in an individual photo. When photos are modified, however, image statistics may be disturbed. For example, if a region is spliced into a photo, an unnaturally sharp edge may be introduced. Forgeries might therefore be detected by estimating statistical properties, and measuring the extent to which they differ from their typical values.

A diversity of natural image statistics can be used to characterize images, and may be disrupted by forgery. For example, statistically-based methods have been developed to detect image splicing. Statistical properties that have been considered include joint probability distributions of colors and gray-level co-occurrence matrices (to characterize texture), statistical properties of wavelet coefficients and multi-size block DCTs, correlations between wavelet coefficients, Markov transition probabilities, phase congruency and bicoherence (to characterize sharp edges), auto-regression coefficients, image quality metrics, and binary similarity measures. Forgery is typically detected by training classifiers to the measure the extent to which such statistics differ from their natural values [160, 90, 248, 56, 223, 36, 87, 222, 11, 78, 88, 247, 5, 181, 67, 61]. Similar statistical features have been used to determine if a photo is computer-generated or photo-realistic [49, 153, 256, 143, 192, 33, 191, 217, 238, 37, 212, 252, 158, 161, 179, 178, 67]. Image statistics can also be used to detect unrealistic color combinations [135].

In addition to natural image statistics, many photos are stored in a compressed format such as JPEG, which introduces distinct statistical properties. The most well studied statistics result when JPEG images are modified and subsequently recompressed in JPEG format. JPEG compression performs a discrete cosine transformation (DCT) on image blocks, and quantizes the resulting DCT coefficients into a smaller range. When JPEG images are recompressed, the already-quantized coefficients are requantized into a new range. This may introduce periodic artifacts into the histogram of the DCT coefficients [206]. These doubly-compressed JPEG images can therefore be identifie by detecting periodicity in the DCT histogram, and measuring when it is more significant than that found in singly-compressed images [39, 149, 168, 197, 92, 206, 202]. Figure 1.3 illustrates this approach. Additional statistical properties have been proposed including periodicity in the image domain [40, 70, 38, 157], and Benford's law in the DCT domain [138, 79]. Small doubly-compressed regions can also be detected [261, 249, 17, 267, 64, 149, 258, 92]. And, many improvements have been developed [15, 16, 40, 74, 6, 97, 47, 148, 209, 224, 154]. JPEG compression statistics have also been used to detect resizing [257, 218] and median filtering [47].

Statistical properties of the image histogram can be used to characterize common photo manipulations. For example, brightening a photo may introduce unnatural gaps or peaks into the histogram. Such changes call into question the originality of a photo. Histogram statistics have been used to detect brightness and contrast adjustments [162, 227, 226, 228, 225], although the evidence can be hidden [28]. The statistical effects of sharpening or blurring have also been considered [27, 24, 47, 145], along with content-adaptive re-scaling (seam carving) [72].

Methods based upon statistical properties are attractive because a diversity of statistics have been studied in prior work, they are often imperceptible to observers, and they are readily combined with classifiers to automate forgery detection. Double JPEG compression

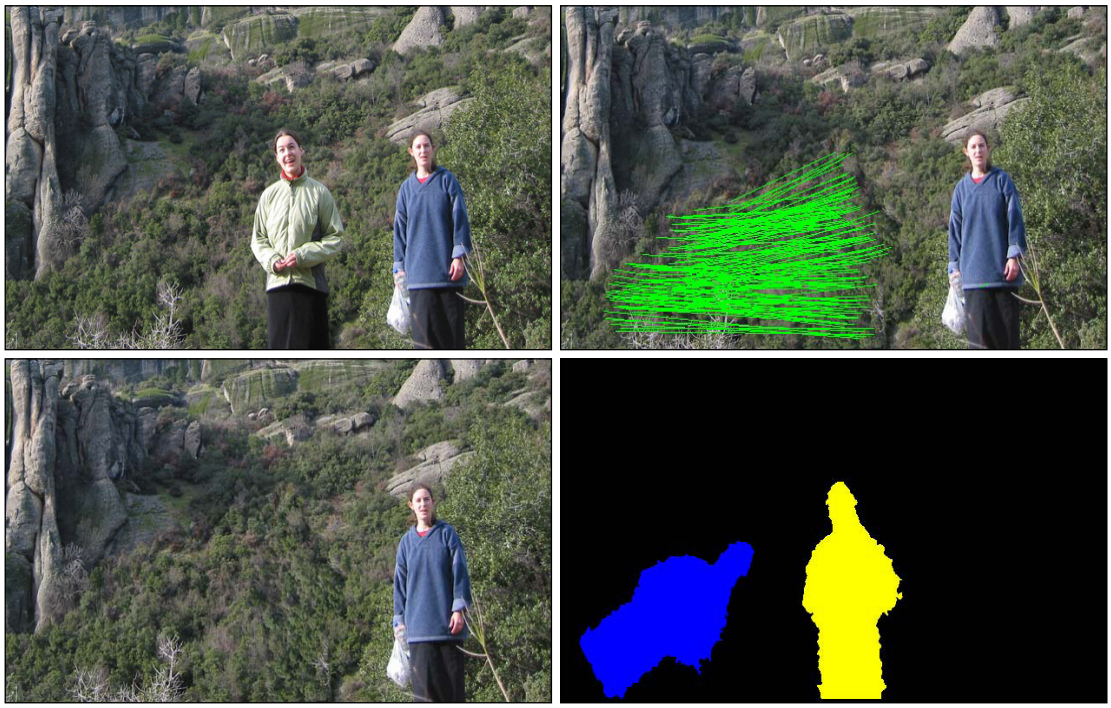


Figure 1.4: Unnaturally similar regions are introduced by forgery when regions are copied and pasted within a photo. These so-called copy-paste forgeries are typically used to hide objects. For example, two people are photographed in an authentic image (top left). Highly regular textures are copied to hide one subject without drawing attention to the manipulation (bottom left). This creates pairs of regions that correspond. The region correspondence property of a photo can be conceptualized as a graph in which edges identify the corresponding regions (top right). The graph shown here has been pruned to show only geometrically regular correspondences, which do not occur naturally. Corresponding regions that are unnaturally similar are retained as evidence of forgery (bottom right). Figure due to Pan and Lyu [193].

can be hidden [229, 231, 230], although this can be detected [244, 245]. Most statistical properties are, however, typically sensitive to resizing and compression, including those that detect double JPEG compression.

Region correspondence

Photos naturally contain regions that are similar, such as patches of grass or sky, which visually correspond to one another. This property can be exploited when falsifying a photo. For example, a region of grass can be copied and pasted elsewhere in a photo to cover an object without drawing an observer's attention to the manipulation. These so-called copy-paste forgeries, may however create an unnatural pattern, and degree, of similarity between regions. Forgeries might therefore be detected by estimating region correspondences, and measuring when they differ from those found in natural photos.

Region correspondence can be viewed as a natural property of photos. Conceptually,

this property is a graph in which edges are drawn between corresponding regions, and are weighted according to the degree of region similarity. In authentic photos, correspondences (edges) are typically directed randomly in the photo, and regions are similar but not identical (e.g., low edge weights). When a photo is modified by copying and pasting a region, correspondences may become geometrically regular, and regions may be nearly identical. Forensic methods have therefore been developed to efficiently estimate region correspondence graphs.

The most common strategy to estimate region correspondence involves dividing a photo into overlapping blocks, compressing the blocks, sorting the compressed blocks to find those that correspond identically, and removing spurious correspondences which are not geometrically regular in some way. The resulting graph is then visually examined to detect forgery [22, 21, 101, 175, 8, 215, 20, 13, 147, 246, 113, 142, 263, 7, 57, 140, 163, 176, 136, 156, 10, 203, 77]. Many of these techniques compute rotation and scale invariant properties from blocks, which improves matching when copied regions have been rotated and warped before pasting. Alternative strategies employ keypoint detectors such as SIFT, rather than blocks [111, 2, 193, 4, 18, 3, 98, 98, 115]. Figure 1.4 illustrates a SIFT-based approach. Comparisons between the various approaches have been performed [46, 44, 45]; both strategies are competitive. Image noise has also been used to estimate region correspondence [196].

Methods based upon region correspondence are attractive, rather than more general forensic approaches, because this unique property can aid forgery detection. Region correspondence is, however, difficult to estimate, and unnatural region correspondence properties can be difficult to define and detect. Such methods may also be sensitive to resizing and compression [184].

Camera response

Photos encode pixel values according to a non-linear mapping from the radiance of the light that strikes the sensor. This mapping is known as the camera response function. Although the response function affects image brightness, it is difficult to visually observe in a photo. When photos are modified, however, the camera response function may be affected. For example, if two images are spliced together, a photo may comprise two distinct camera response functions. Forgeries might therefore be detected by estimating camera response properties, and measuring the extent to which they differ within a photo.

The camera response function of a photo is typically monotonically increasing and smooth, and it is straightforward to estimate if both the image brightness and original radiance values are known. Radiance is however rarely known. Camera response may nonetheless be estimated from pixel intensity if the form of the response function is constrained. Methods have therefore been proposed to estimate camera response in a photo, and differences between estimates can be used to detect forgery. Such methods have used

histogram properties [25], information from points in the image where the change in radiance is approximately linear [96, 183, 95, 182, 94], information from edges [150], and high-order signal correlations (bicoherence) [206, 202].

Methods based upon camera response are attractive because cameras typically employ unknown response functions, and are therefore rarely corrected by a naive forger. There is also evidence that estimation of camera response can be robust JPEG compression [96]. These methods, however, may suffer if cameras employ similar response functions. Camera response is also a global manipulation, and might therefore be easily corrected.

Metadata

Photos are encoded with metadata properties that describe camera settings which were used during capture. Much of this metadata can be arbitrarily chosen by the camera manufacturer, and the combination of many such arbitrary choices may be distinct to the camera make and model. When photos are modified, however, metadata is typically changed by photo-editing software. Modified photos might therefore be detected by extracting metadata properties, and measuring the extent to which they differ from the metadata that is produced by known cameras.

A diversity of metadata properties may be embedded in the header of digital photos. Methods that use metadata for forensic purposes build databases of the property values that are employed by digital cameras of differing makes and models. Such databases can be built by downloading original photos from internet photo-sharing websites. Useful metadata properties have included JPEG compression parameters (quantization tables and Huffman codes for each color channel), parameters of the embedded JPEG "thumbnail" image (quantization tables and Huffman codes), as well as the number of additional discretionary data elements (EXIF fields such as GPS data) [118, 116, 63, 134, 62, 1]. These metadata properties have been shown to be highly distinct: databases have been built which comprise thousands of cameras.

Methods based upon metadata are attractive because such properties are easily estimated from a photo — the appropriate fields are simply extracted from the header. Furthermore, large numbers of photos can be quickly scanned. These properties, however, only help to determine if a photo is unmodified, rather than to detect tampering. Metadata can also be modified without impacting image content, and this aids forgery.

Summary

Digitally-based forensic methods have considered the properties of noise, pixel correlation, image statistics, region correspondence, camera response, and metadata. These properties are useful for forensics because they are present in most photos, their estimation is typically well-posed, their computation can be automated for the analysis of many photos, and they

are typically visually imperceptible.

While the visual subtlety of digital properties is important, methods are typically sensitive to resizing and compression, which destroys such subtle visual information. Sensitivity varies among methods, but a rule of thumb might be stated that downsizing by 50% and compressing at 75% JPEG quality will destroy the currently-employed digital properties. This nearly universal attack is troubling because it is naively employed when photos are distributed on the internet; resizing a typical 5 MP or 10 MP photo to the relatively large web resolution of 1 MP constitutes resizing factors of 50% and 30%. Furthermore, small image size does not evidence manipulation, as most cameras can be natively configured to produce small images. Another challenge results because many digital properties can be automatically estimated. While automation facilitates the analysis of many photos, it also aids forgery; many digital properties can be quickly adjusted by a forger.

Partially in response to these challenges, forensic methods have been proposed that measure physical properties in a photo, which result from the camera optics or scene. Many of these physically-based methods are robust to resizing and compression, and require that an informed forger perform tedious manual photo modifications. Physically-based methods, of course, present their own challenges.

1.1.2 Physical properties

Before photos are digitized, light undergoes a series of physical transformations which ultimately result in the formation of an analog image on the camera sensor. These physical processes give photos physically-based properties that are often visually subtle, and may be inadvertently modified by forgery. A diversity of forensic methods have therefore been proposed to estimate physical properties, and to measure the extent to which they differ from their appropriate values.

Perspective

Photos can be described as a geometric projection of points in the 3-D world onto a 2-D image plane, according to a linear perspective transformation. The effect of varying the perspective of a photo is typically visibly identifiable, but perceptually subtle. When a photo is modified, however, its perspective properties may be inadvertently changed. For example, an object may be moved within a photo, but this cannot change the angle from which the object was originally photographed. This effect can be described by properties of perspective projection. Forgeries might therefore be detected by estimating perspective properties, and measuring the extent to which they differ from their typical values.

Perspective projection properties are embodied by the so-called camera matrix, which defines the intrinsic properties of the camera, such as its focal length, and its extrinsic properties, such as its location in the world. Estimation of projection properties typically

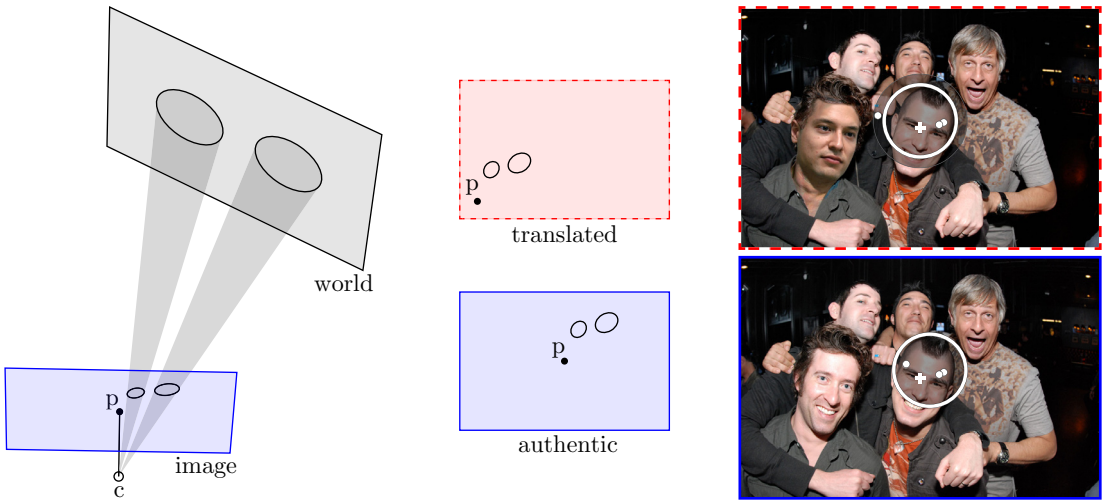


Figure 1.5: The principal point of the camera, a property of perspective projection, is modified when regions are translated within a photo. **Left:** pairs of circles, which project into the image plane (blue) as ellipses, can be used to estimate the principal point, p, of a camera (this camera has center of projection c). The principal point typically lies near the image center. **Center:** if these circles are translated in a photo (dashed red) the principal point is implicitly moved. **Right:** pairs of circles occur in people's eyes, and the principal point can be estimated for each pair. Shown is a fake image (top) in which a person's face was replaced (original photo at bottom). Top right: the estimated principal point (white dot) for the modified face deviates too far from image center (white circle); forgery is detected. Bottom right: in the unmodified photo all principal point estimates lie close to the image center (within the white circle). Figure due to Johnson and Farid [109].

requires some knowledge of the geometry of objects in a photo. For example, if people are present, the known geometry of their eyes (pairs of circles) can be used to estimate the optical center of the image [109]. The optical center, or principal point, should be close to the image center, but it is moved if subjects are moved within the photo. Deviations of the principal point therefore evidence forgery. Figure 1.5 illustrates this approach.

Similarly, deviations of the intrinsic skew parameter of a camera can be measured to detect bootleg video recordings; skew is introduced if videos are recorded from an oblique angle in a theatre [250]. If signs or billboards are present in a photo, their text may be falsified such that it is inconsistent with proper perspective projection. The known geometry of fonts can be used to estimate the camera matrix, and to detect forgery [48]. Projective constraints which entail pairs of purportedly authentic images of the same scene have also been considered [264]. Lastly, reflections of objects in windows and other flat surfaces can be used to estimate and detect inconsistent perspective [186] (more on reflections below).

Methods based upon perspective properties are attractive because perspective errors are difficult to perceptually subtle. Perspective is also challenging to correct during forgery because 2-D image editing software cannot emulate changes in 3-D perspective. These methods, however, require geometric constraints, which may not be available in a photo.

Optics

Photos are formed by focusing light through a series of optical elements, and onto an optical sensor. A variety of visible optical properties are produced by this process, but may be misunderstood by observers. When photos are modified, however, optical properties may be inadvertently changed. For example, if an object is moved within a photo, it may be placed into a region of the photo that contains differing geometric distortions. Forgery might therefore be detected by estimating optical properties, and measuring the extent to which they differ from their typical values.

A variety of optical properties have been considered. Chromatic aberration, which manifests as unnatural colors along edges, is often visible in photos but may appear somewhat random to an observer. These artifacts however form a geometric pattern that can be disrupted by forgery. For example, if a region is moved within a photo, a radial pattern formed by lateral chromatic distortions may be disrupted. Forgery might therefore be detected by estimating the lateral chromatic aberration in a photo, and measuring when this property differs between regions [106]. Methods have also considered longitudinal chromatic aberration and purple fringing [259].

Light that passes through an optical system is attenuated before it reaches the sensor, and the extent of this attenuation may vary across an image to form a so-called vignetting pattern. Vignetting typically produces darkened corners in a photo, which are perceptually subtle but visually identifiable. When photos are modified, vignetting may be disrupted; for example, cropping a photo will shift this pattern. Forgeries might therefore be detected by estimating vignetting, and measuring when it is atypical. For example, vignetting can be used to identify the lens model [159], which might be compared to the image metadata.

Optical systems also create blur, which results from both lens imperfections and optical design. Defocus blur, which is often intentional in photography, should be constant if objects have the same depth. When photos are modified, however, regions may be inserted with differing blur but apparently similar depths. Forgery might therefore be detected by estimating defocus blur, and detecting regions that have different amounts of blur [251, 237]. Inconsistent motion blur has also been considered [112].

Lastly, commercial lenses are typically designed to produce linear perspective images, but non-linear distortions are often present. These geometric lens distortions warp straight lines, and may be visually observable. The pattern of geometric distortion is however less obvious. When photos are falsified, for example by moving or splicing a region, distorted edges may violate what is typically a radial pattern of distortion. Forgery might therefore be detected by estimating lens distortion, and measuring where it differs within a photo [41].

Methods based upon optics are attractive because such properties are often visually observable, preserved by resizing and compression, and may be perceptually subtle to a forger. It may however be difficult to estimate properties if information is limited. For

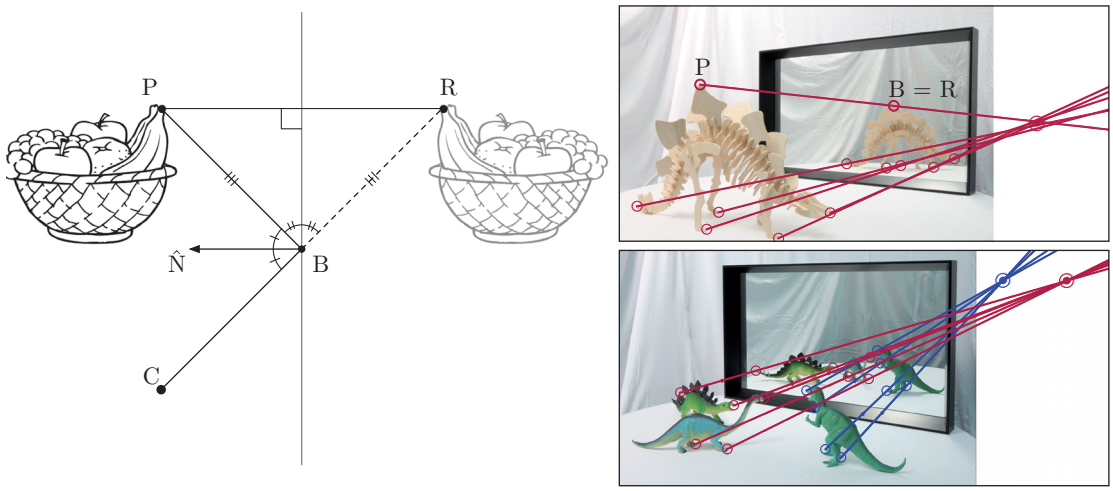


Figure 1.6: Inconsistent reflection properties may be introduced when photos are modified. A simple line construction in the image can detect inconsistent reflections. Left: the geometry of reflections. Point P on an object is reflected at point B from the perspective of a camera located at point C. Top right: an observer identifies points P and B in an image and draws the line PB. Because point R (shown in left panel) also lies along the camera ray through C and B, the projected location of R and B are the same (top right). The line PB therefore identifies the image of a ray PR that is perpendicular to the reflecting surface (also shown in left panel). All rays PR are parallel, and must therefore converge to the vanishing point (top right) of the surface normal direction $\hat{N}$. Bottom right: multiple vanishing points evidence forgery. Figure due to O'Brien and Farid [186].

example, lens distortion requires lines that are known to be straight in the world.

Reflection

Photos may contain reflections of objects in windows, water, and other surfaces. Although reflections may be visually salient, their physical consistency is often perceptually subtle. When photos are modified, however, the physical properties of reflections may be violated. For example, if an object is moved within a photo, its reflection in surfaces may not appear in the correct location. Forgeries might therefore be detected by estimating reflection properties, and measuring the extent to which they differ from their typical values.

Reflection properties can be used to estimate what is, more precisely, a geometric property of the scene. Specifically, it can be shown that lines which connect points on an object to their reflection must converge to a point.[1] This convergence point is the vanishing point of the vector that is normal to the reflecting surface [186], and may be perturbed when reflections are falsified. For example, if an object is rotated but its reflection is not correctly adjusted, lines may not converge to a point. Forgery can therefore be detected by manually identifying correspondences between an object and its reflection, and constructing lines; differing vanishing points evidence forgery [186]. Figure 1.6 illustrates this approach.

[1] This point may lie at infinity in the image plane.

Figure 1.7: Lighting properties may be perturbed when photos are modified. **Left:** the brightness of a patch $\vec{x}$ on a Lambertian surface depends upon the amount of light that strikes the patch from all directions $\vec{V}$ on a hemisphere centered about the surface normal $\vec{N}$ of the patch. The lighting can therefore be described as a function over the surface of a hemisphere. **Middle:** a falsified photo. **Right:** surface normals on the occluding contour (white) of objects can be used to estimate 2-D components of the lighting function. Shown is the lighting for the subject's head and body (spheres 1 and 2). Differences in lighting evidence forgery. Figure due to Johnson and Farid [107].

By a similar approach, reflections can be used to estimate the projected location of an object that is reflected in multiple surfaces of known orientation. For example, the location of a point light source can be estimated when it is reflected in the eyes of multiple people. The known geometry of the human eye can be used to estimate the surface normal at each reflection, which projects to a vanishing point in the image. Each vanishing point must lie on a line in the image that passes through the reflection, and through the light source. This relationship is visually subtle, and may be perturbed if subjects are moved (or spliced) in a photo. Forgery can therefore be detected by using the reflections in peoples' eyes to estimate the location of the light source, and detecting differing light positions. This approach has been equivalently described in the literature in terms of estimating light direction [216, 108].

Methods based upon reflection are attractive because reflections are often visually salient, and therefore preserved by resizing and compression. Furthermore, observers have difficulty judging, and falsifying, consistent reflections. Manual input is however required to identify corresponding points, and reflections may not be present in a photo.

Lighting

Photos capture the interaction of light and the 3-D surfaces in a scene. Although lighting is visually salient in photos (in the form of shading and shadows), inconsistencies in lighting may be perceptually subtle. When photos are modified, however, lighting properties may be changed. For example, objects that are spliced into a photo may have originally been photographed under different lighting conditions than are possible in the photo. Forgeries might therefore be detected by estimating lighting properties, and measuring the extent to which they differ from their appropriate values.

Lighting properties manifest in part as the shading on diffuse objects, and are therefore perturbed when objects are inserted or moved within a photo. Although light estimation typically requires knowledge of 3-D object geometry [9, 210], 2-D estimates can be computed from the occluding contours of objects [185]. Forgery can therefore be detected by estimating the light direction or environment, and measuring differences between the estimates for various objects in the photo [117, 236, 107, 105]. Figure 1.7 illustrates this approach.

Lighting properties also manifest in photos as shadows, which can be used to estimate the projected location of a point light source. Point light sources must lie on a line that connects a point in shadow to the corresponding point on an object. Furthermore, if multiple shadows are present, all lines must intersect at the light source. Inconsistent shadows may however seem plausible to an observer [66, 190, 103], and may be created if objects are moved in a photo. Forgery might therefore be detected by identifying lines between corresponding points on shadows and objects, and measuring differences between the intersection points of the lines. This method has been employed for the analysis of shadows in paintings [233, 234, 233, 232, 172]; a similar analysis might also be used in photos. Other line constructions have been proposed that use shadows, but assume that the ground plane is flat and orthogonal to the object [265].

Methods based upon lighting are attractive because inconsistencies may be overlooked by observers, while lighting often manifests as low-frequency signals that are preserved by resizing and compression. Lighting properties are also difficult to modify, as this typically requires manual photo adjustment. Methods based upon lighting, however, require geometric constraints, and typically restrict the lighting conditions that can be analyzed.

Summary

Physically-based forensic methods have considered properties of linear perspective, optics, reflection, and lighting. The strengths and weaknesses of these methods contrast those that use digital properties. While digital properties might be automatically modified and corrected during forgery, physical properties often require tedious manual modifications. While digital properties may be destroyed by resizing and compression, many physical properties manifest as low frequency, visually observable, signals that are preserved.

Conversely, estimating physical properties often requires manual input, while digital properties can be automatically estimated to analyze many photos. And while physical properties may restrict the types of scenes that can be analyzed, digital properties afford analyses that are comparatively independent of scene content. Digital and physical properties therefore fulfill somewhat complimentary roles. Both types of methods may however be applied to a particular photo. Although no single tool can detect every forgery, the combination of many tools make photo forgery increasingly difficult and time consuming.

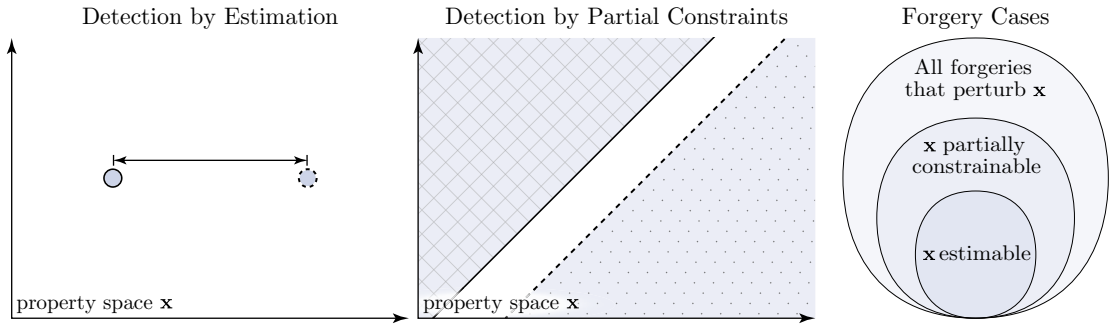


Figure 1.8: Forgery detection by estimation, and by partial constraints. Both approaches detect when a property $\mathbf{x}$ of a photo is perturbed. The estimation approach (left) computes multiple estimates of $\mathbf{x}$ (e.g., solid and dashed dots) from regions of a photo, and differences between these estimates evidence forgery (measure line).[2] The partial constraint approach (center) relaxes the estimation requirement by constraining the range of plausible values, $\mathbf{x}$ (e.g., solid and dashed regions), and computing constraint satisfiability (regions should intersect). Partial constraints may extend to infinity in one or more dimensions. By eschewing estimation, forensic methods can be generalized to detect forgery in photos for which the required property cannot be estimated (right).

1.2 Partial constraints

A large and growing body of forensic methods have been developed to detect forgery, each targeting a property that may be perturbed when a photo is modified. Recent physically-based methods have however brought into focus a common challenge in forensics: it is often difficult or impossible to estimate properties from regions within an image of unknown and uncontrolled origin. This difficulty has restricted the application of forensic methods, which assume that explicit estimation of properties is necessary.

We present an alternative approach in which forgery is detected by partially constraining a property, rather than explicitly estimating it. For example, it may be impossible to estimate the direction of a light source in a particular photo, but it might be straightforward to partially constrain that direction to one side of the photo. Photographs are viewed as a collection of many partial constraints, each quantifying the information that a region of the photo provides about the property. Although each individual constraint may provide limited information, the collection of many constraints can be highly informative. If the information in a photo specifies satisfiable constraints, the photo can be determined to be plausible. Conversely, if a photo contains conflicting information, as evidenced by unsatisfiable constraints, it can be determined to be falsified. By relaxing the requirement that properties be explicitly estimated, forensic analyses can be generalized to a broader range of photos, Figure 1.8.

[2]Methods that compare a single property estimate to a known value that is external to a photo can be viewed as making comparisons to a default property estimate that is present in all photos. Externally defined thresholds or classification boundaries may appear similar to a partial constraint, but do not relax the requirement that properties be estimated in a photo.

1.3 Contributions

We develop complimentary physically-based methods for the analysis of shadows and shading in photos for which explicit estimation of the associated lighting properties is not possible, Chapter 2 and Chapter 3. We then describe a variety of prior methods that can be generalized to detect forgery in photos for which it is not possible to explicitly estimate the required properties, Chapter 4. We conclude in Chapter 5.

In Chapter 2, shadows are used to partially constrain the 2-D projected location of a light source, thereby enabling detection of inconsistent shadows in common settings such as outdoor scenes. We describe how ambiguously shaped cast and attached shadows can be used when they form under both point and area lights. In contrast, prior methods have been limited to the analysis of photos that contain well-defined cast shadows formed by point lights [110, 234, 233, 232].

In Chapter 3, the shading on objects is used to partially constrain the 2-D projected location of a point light source, thereby enabling detection of inconsistent combinations of shading and shadow. We describe how shading and shadow cues can be combined in photos that are captured under unknown linear perspective, and that contain objects for which their 3-D geometry can be only partially constrained. In contrast, prior methods have been limited to the analysis of paintings for which the assumption of orthographic projection is sufficient [110, 234, 233, 105].

In Chapter 4, a variety of prior forensic methods are reformulated in terms of partial constraints, and we show that each can be generalized to detect forgery in photos for which estimation is not possible. Specifically, we reformulate methods based upon the camera center of projection, optical chromatic aberration, and specular reflections.

In Chapter 5, we conclude with a brief discussion of the limitations of partial constraints and new questions that are raised by the approach. By expanding forensic analyses to include photos for which estimation is not possible, it is our hope that new forensic methods can be designed to detect forgery in a broader range of photos. Although no method can make it impossible to craft a convincing forgery, forensic methods can however require that photographs tell a digitally and physically consistent story by detecting inconsistencies in the information that is present.

Chapter 2

Shadows

In this chapter, we describe a method for determining if cast and attached shadows in a photo are consistent with the model of a single distant or local point light source. We then extend this method to include the analysis of scenes which are illuminated by a single area light. Forensic techniques based on analyzing lighting and shadows are attractive because 3-D lighting effects can be difficult to modify using commercial photo editing software, and low quality images can be analyzed since lighting effects and shadows survive common operations such as image compression and down-sizing.

There is some evidence that the visual system is capable of detecting small changes in lighting direction in simple controlled settings [133, 172, 119, 201, 131, 189]. In more complex settings, however, the visual system is far less capable at detecting gross inconsistencies in lighting [103, 190, 66]. In a forensic setting, a multitude and variety of cast and attached shadows from complex shapes are cast onto equally complex and varied surfaces. Such limitations of the visual system imply that a forger may overlook inconsistencies in lighting and shadow, and a visual inspection of shadows will, at best, be highly subjective.

The iconic photo of the 1969 moon landing, Figure 2.1, provides an example of the complexity of shadows that are common in photos. In fact, it has been argued by conspiracy theorists that the shadows in this photo are physically implausible and hence evidence of photo tampering and broader nefarious conspiracies. Beyond a subjective visual analysis, the physical consistency of shadows can be determined by considering their basic geometry.

Consider a ray that connects a point in a shadowed region to its corresponding point on the shadow-casting object. In the 3-D scene, this ray intersects the light source. In a 2-D image of the scene created under linear perspective, the projection of this ray remains a straight line that must connect the images of the shadow point and object point, and intersect the projected image of the light source. These constraints hold regardless of the geometry of the object and the surface onto which the shadows are cast, and for either an infinitely distant or local light. Multiple constraints can therefore be used to determine the projected location of a light source in the image plane. Note that this projected location corresponds to an infinite number of 3-D light positions. For simplicity we concern ourselves

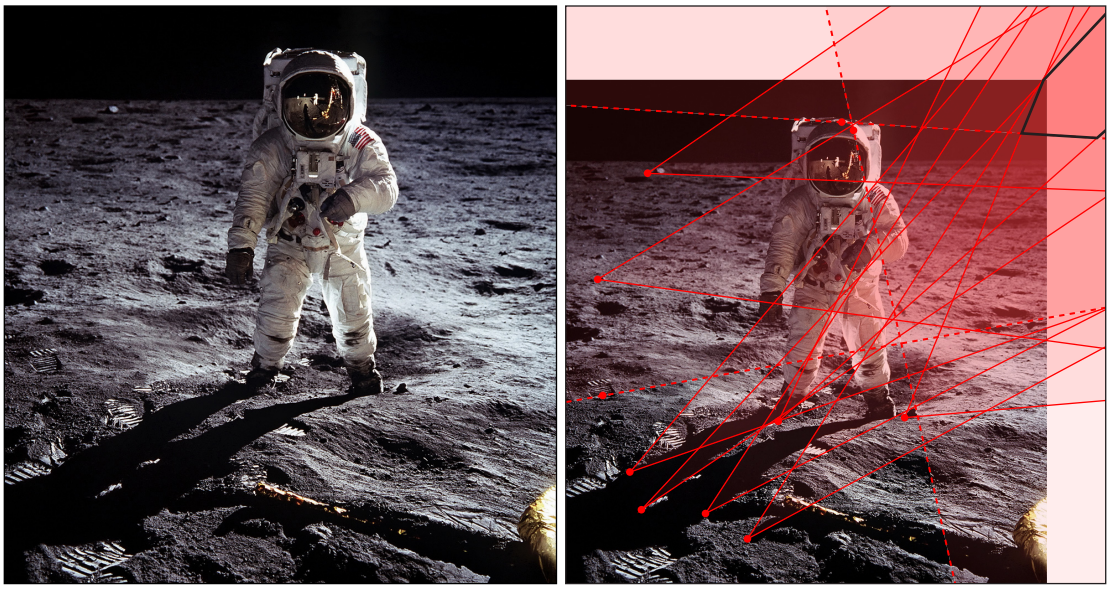


Figure 2.1: Our algorithm finds that the shadows in this 1969 moon landing photo are physically consistent with a single light source. The solid lines correspond to constraints from cast shadows and dashed lines correspond to constraints from attached shadows. The region outlined in black, which extends beyond the figure boundary, contains the projected light locations that satisfy all of these constraints.

with the 2-D projection because the 3-D location of a light source is typically highly under-constrained in a photo.

If a scene purportedly contains a single light source but the shadows in the scene specify mutually inconsistent constraints that cannot be satisfied by any single light position, then this inconsistency evidences photo tampering. It can be difficult or impossible, however, to precisely match a point on a shadow to its corresponding point on an object — particularly on attached shadows that form as a surface smoothly curves to face away from the light.

We therefore consider a relaxed, conservative, partial constraint in which the location of points on the object are restricted to a range of possible locations. These partial constraints specify either angular wedges or half-planes in the image that restrict the projected location of the light source. (See Figure 2.2.) The satisfiability of multiple such constraints is framed as a linear programming problem. A satisfiable set of constraints is interpreted to mean that the shadows are physically plausible, while unsatisfiable constraints are used as evidence of photo tampering.

Related Work

Previous lighting-based forensics methods estimate the 2-D lighting direction or lighting environment from the shading on an object's contour [105, 107]. If the 3-D geometry of an object is known, then the 3-D lighting direction or lighting environment can be estimated [117]. Related computer vision techniques that estimate lighting from a single image

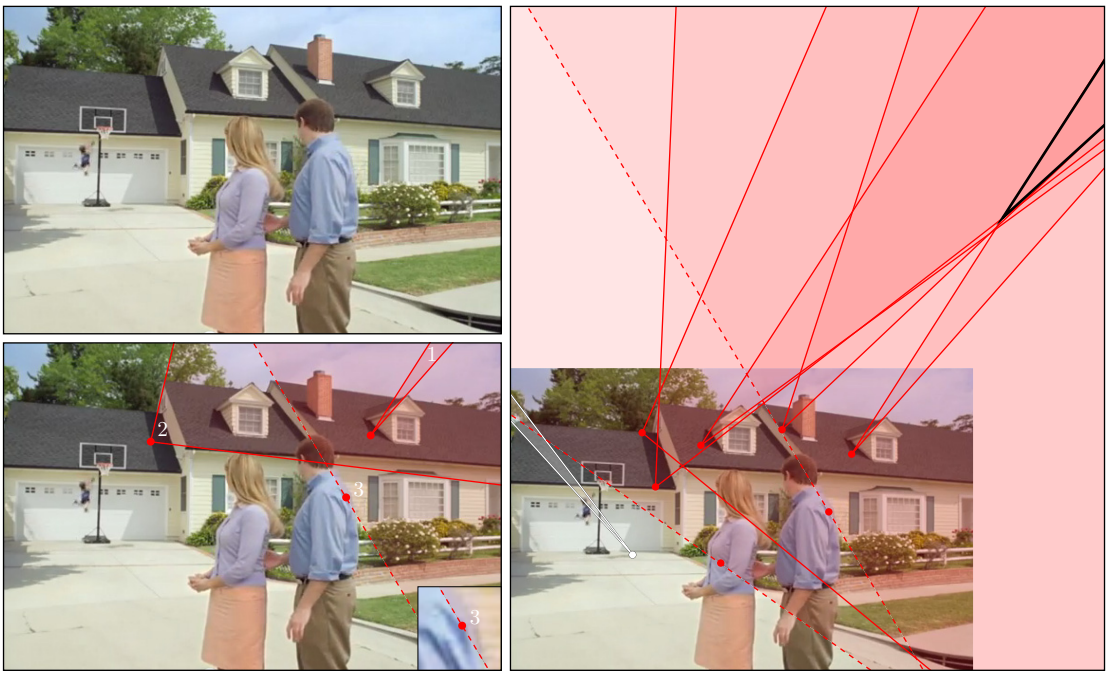


Figure 2.2: Shown are: (top left) a frame from the Geico commercial "Dunk – Easier Way to Save" depicting a somewhat incredible athletic performance; (bottom left) examples of cast (1-2) and attached (3) shadow constraints. The projected location of the light source must lie in the intersection of these constraints. Shown on the right are a variety of constraints (solid lines correspond to cast shadows and dashed lines correspond to attached shadows). The shadows from the people and house are consistent with a light source located somewhere in the black-outlined region. The boy's shadow, however, is inconsistent with the rest of the scene (white shaded constraint).

use object shading [185], or shadows cast onto planar surfaces [219, 187, 135]. Manually approximated scene geometry has been used to fit a local lighting model that is perceptually plausible, but insufficient for forensic application because the physical accuracy is heavily influenced by user input [114].

Photometric inconsistencies of a cast shadow's umbra have been used to detect inconsistent shadows [151], under the assumption of a ground plane that has constant reflectance. Inconsistencies in the location of a cast shadow have also been considered [265], but several assumptions were placed on the scene geometry: shadows were cast onto a planar ground plane and the objects casting shadows were vertical relative to the ground plane. In the most closely related work [234], the consistency of cast shadows in artworks was determined by identifying points on a cast shadow and their corresponding point on an object, and then determining if these were consistent with a single light source. Our work expands this basic idea by relaxing the requirement that a strict shadow-to-object correspondence must be explicitly estimated. We therefore allow a broader range of ambiguous cast shadows to be considered, including attached shadows. Unlike some previous shadow-based forensic techniques, we place no assumptions on the scene geometry.

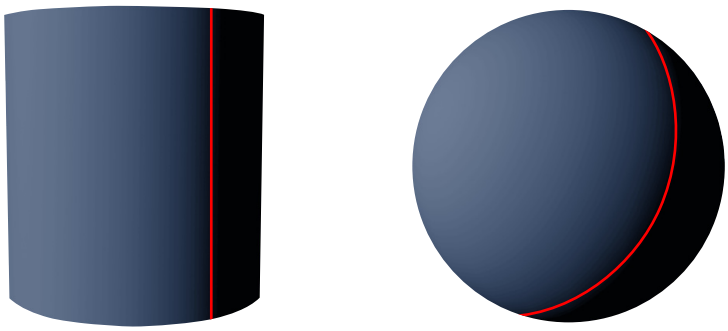


Figure 2.3: A cylinder and sphere are illuminated from the left with a distant point light. The red line and curve identify the terminator — the contour where the surface normal is oriented 90 degrees from the direction to the light source. Points to the right of the terminator are in an attached shadow.

2.1 Methods

We first describe the geometry of cast and attached shadows and how these shadows partially constrain the projected location of a light source. Throughout, we assume a single distant or local light source and place no assumptions on the objects being illuminated or the surfaces onto which shadows are cast. We then frame the problem of determining if all shadows are consistent as a linear programming problem. In the case when the shadows are not consistent, we describe a simple randomized algorithm for finding an approximately minimal set of conflicting constraints that identifies the inconsistent shadows.

2.1.1 Cast shadows

Shown in Figure 2.2 is a frame from a Geico commercial with several shadows. Wedge-shaped constraints are used to describe the directions from a point in shadow to points on an object that may have cast the shadow. The wedge labeled 1 corresponds to a cast shadow on the roof. This wedge is fairly narrow because the tip of the cast shadow can reliably be determined to correspond to a point on the dormer. The wedge labeled 2 corresponds to a cast shadow on the garage roof. This second constraint is a wider wedge because the correspondence between the roof edge and its cast shadow is ambiguous.

The projected location of the light source should lie within the intersection of these wedges (which although illustrated as finite, are infinite in their extent beyond the figure boundary). Note that the wedges are oriented from the shadow towards the corresponding object. If, however, the light is behind the camera then these wedge constraints should be flipped 180 degrees about the selected shadow point. That is, due to perspective geometry there is a sign ambiguity as to the location of the projected light source (more on this issue below).

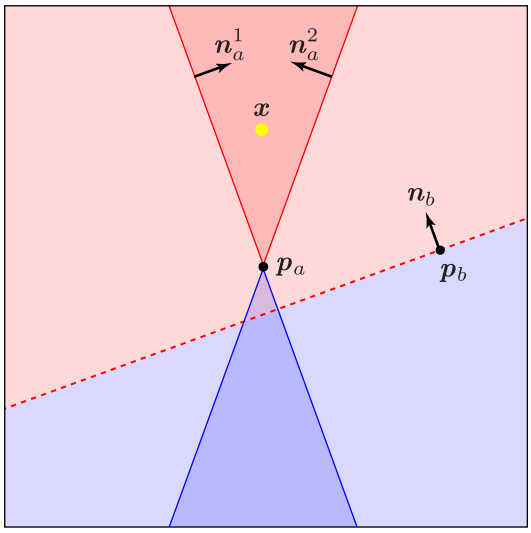


Figure 2.4: A wedge constraint defined by two lines (solid red) and a half-plane constraint defined by one line (dashed red). The projected location of the light source (yellow dot, $\boldsymbol{x}$) lies within the region formed by the intersection of these constraints. A sign inversion of these constraints is depicted in blue, and corresponds to the case in which the location of light sources behind the camera are inverted by projection.

2.1.2 Attached shadows

Attached shadows occur when objects occlude the light from themselves, leaving a portion of the object in shadow. For example, shown in Figure 2.3 is a cylinder and sphere illuminated by a distant point light source positioned to the left. Points are in shadow if the surface normal makes an angle greater than 90 degrees with the direction toward the light. The terminator is defined to be the surface contour whose normals form a 90 degree angle with the direction toward the light source, as depicted by the red line and curve in Figure 2.3.

Similar to a cast shadow, there is a correspondence between points in and out of shadow on either side of the terminator. This correspondence, however, can only be specified to within a half-plane because the light's elevation is ambiguous to within 180 degrees.

Although cylinders and spheres may not be particularly common in natural photographs, any locally convex surface can provide an attached shadow half-plane constraint. Shown in Figure 2.2, for example, is an attached shadow constraint on a fold of the man's shirt (constraint 3 in the middle panel). Folds and other locally convex geometry are common and provide easily recognizable attached shadow constraints.

2.1.3 Forensics from shadows

For an authentic image there must be a location in the infinite plane ($\mathbb{R}^2$) that satisfies all cast and attached shadow constraints. That is, the intersection of all the constraints should define a non-empty region. In the right panel of Figure 2.2, for example, there is a non-empty intersection for all of the shadow constraints on the parents and house — the region

outlined in black. Note, however, that the shadow of the airborne boy in the background generates a wedge constraint (shaded white) that does not intersect the other constraints and is therefore physically inconsistent for a scene with one light source.

Cast and attached shadow constraints can be represented as linear inequalities in the plane. The satisfiability of these constraints can then be determined using standard linear programming. Shown in Figure 2.4 are two lines defined implicitly by their normals $\boldsymbol{n}_a^1$ and $\boldsymbol{n}_a^2$, and the point $\boldsymbol{p}_a$. The direction of the normals specifies the region in the plane in which the solution $\boldsymbol{x}$ must lie. The intersection of these two regions is the upward facing wedge (red). Also shown in Figure 2.4 is a single line defined by its normal $\boldsymbol{n}_b$ and point $\boldsymbol{p}_b$. This line specifies a half plane constraint in which the solution must lie. In each case, a shadow constraint is specified by either a pair of lines (wedge cast shadow constraint) or a single line (half-plane attached shadow constraint).

Formally, a half-plane constraint is specified with a single linear inequality in the unknown $\boldsymbol{x}$:

$$\boldsymbol{n}_i \cdot \boldsymbol{x} - \boldsymbol{n}_i \cdot \boldsymbol{p}_i \geq 0, \tag{2.1}$$

where $\boldsymbol{n}_i$ is normal to the line and $\boldsymbol{p}_i$ is a point on the line. A wedge-shaped constraint is specified with two linear constraints:

$$\boldsymbol{n}_i^1 \cdot \boldsymbol{x} - \boldsymbol{n}_i^1 \cdot \boldsymbol{p}_i \geq 0 \quad \text{and} \quad \boldsymbol{n}_i^2 \cdot \boldsymbol{x} - \boldsymbol{n}_i^2 \cdot \boldsymbol{p}_i \geq 0. \tag{2.2}$$

A collection of half-plane and wedge constrains can be combined into a single system of m inequalities:

$$\begin{pmatrix} \boldsymbol{n}_1 \\ \boldsymbol{n}_2 \\ \vdots \\ \boldsymbol{n}_m \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} \boldsymbol{n}_1 \cdot \boldsymbol{p}_1 \\ \boldsymbol{n}_2 \cdot \boldsymbol{p}_2 \\ \vdots \\ \boldsymbol{n}_m \cdot \boldsymbol{p}_m \end{pmatrix} \geq \mathbf{0} \tag{2.3}$$

$$\mathbf{N}\boldsymbol{x} - \mathbf{P} \geq \mathbf{0}. \tag{2.4}$$

Given error-free constraints from a consistent scene, a solution to this system of inequalities should always exist. However we can account for errors or inconsistency by introducing a set of m slack variables s_i:

$$\mathbf{N}\mathbf{x} - \mathbf{P} \geq -\mathbf{s} \tag{2.5}$$

$$\mathbf{s} \geq \mathbf{0}, \tag{2.6}$$

where s_i is the i^{th} component of the m-vector $\mathbf{s}$. If the constraints are fully satisfiable then a solution will exist where all slack variables are zero. Solutions with a non-zero slack variable

s_i mean that the solution is not consistent with constraint i.

The above inequalities can be combined into a single system as follows:

$$\begin{pmatrix} \mathbf{N} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ \mathbf{s} \end{pmatrix} - \begin{pmatrix} \mathbf{P} \\ \mathbf{0} \end{pmatrix} \geq \mathbf{0}, \tag{2.7}$$

where $\mathbf{I}$ is an $m \times m$ identity matrix. We, of course, seek a solution that minimizes the amount of slack required to satisfy this system. As such, the linear program consists of minimizing the L_1 norm of the vector of slack variables,

$$\begin{pmatrix} \mathbf{0} & \mathbf{1} \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ \mathbf{s} \end{pmatrix}, \tag{2.8}$$

subject to the constraints in Equation (2.7). That is, we seek a solution that minimizes the slack variables, while satisfying all of the cast and attached shadow constraints. If the slack variables for the optimal solution are all zero, then there exists a light position that satisfies all of the specified constraints. Otherwise, one or more of the shadows is inconsistent with the rest of the scene.

Recall that a shadow constraint is specified by connecting a point on a shadow with a range of possible corresponding points on an object. As described earlier there is an inherent sign ambiguity in specifying these constraints: if the light is behind the center of projection, then its projected location in the image plane is inverted and the constraint normals should all be negated. In this case, the constraints simply take the form:

$$-\mathbf{N}\mathbf{x} + \mathbf{P} \geq -\mathbf{s}. \tag{2.9}$$

In Figure 2.4, for example, the downwards facing region shaded blue corresponds to a sign inversion of the upward-facing region shaded red. In practice we solve both linear programs (with constraints Equation (2.5) and (2.9)) and select the solution with the minimal L_1 norm, Equation (2.8). If either the regular or inverted system has a solution with zero slack, then we conclude that the constraints are mutually consistent. Otherwise, there is no light position that is consistent with all of the constraints and we conclude that some of the constraints are being generated by parts of the image that have been manipulated. Lastly, note that an analyst may wish to specify constraints from an object to its shadow, rather than from a shadow to its object point. In this case, their constraint is imply inverted before constructing Equation (2.5).

When an image generates inconsistent constraints, we may wish to know which constraints are in conflict with others. These conflicting constraints provide the essential evidence that can be used to invalidate a forgery, and can be useful in determining what parts of an image may have been manipulated. We greedily find an approximately minimal set

of inconsistent constraints.[1] To begin, two constraints are selected at random. If these constraints are not satisfiable, then they form a minimal set of inconsistent constraints. If they are satisfiable, then a randomly selected constraint is added to the set and the linear program is solved. Constraints are added in this way until the system is no longer satisfiable. This entire process is repeated with different random starting conditions. The smallest set of violating constraints, which may or may not be unique, provides a succinct summary of which parts of an image may have been altered.

2.1.4 User interface

A multitude of shadows may be present in a photo, each presenting its own multitude of possible constraints that an analyst may specify. This presents a visual search problem, and the successful detection of a forgery depends upon an analyst's ability to identify conflicting constraints. A user interface was developed that guides an analyst to efficiently identify a concise set of constraints that may detect forgery.

A forensic analyst's specification of cast shadow constraints is made easier and more efficient by an interactive process of constraint entry, Figure 2.5. Cast shadow constraints are specified by selecting a point in shadow and selecting two additional points which define the edges of a wedge (red point and lines) that comfortably encompasses the corresponding object casting the shadow, Figure 2.5(a).

After the first constraint is entered, a yellow wedge-shaped region is automatically rendered and follows the user's cursor, Figure 2.5(b). This yellow region is the smallest wedge, emanating from the current point, that would completely cover both the forward and backward feasible regions of the linear programs associated with the previously entered constraint(s). Intuitively, this yellow region identifies for an analyst the directions from the cursor in which she should expect to find an object that has cast a shadow, given the constraints that have been previously entered.

The yellow region guides the analyst's search for additional, useful, constraints. Specifically, the analyst's goal is to reduce the size of this yellow region by identifying new constraints that more tightly encompass the shadow-casting object, Figure 2.5(c). Each such additional constraint implicitly reduces the region of possible locations of the light source, and therefore more precisely identifies possible shadow-to-object correspondences in the photo.

If the wedge that the analyst would need to specify in order to encompass the shadow-casting object fully encloses the yellow wedge, then the new constraint is redundant with previous ones and can simply be ignored, Figure 2.5(d). The yellow wedge therefore allows an analyst to quickly scan a photo for useful constraints (those which further reduce the possible locations of the light source). If during this search the analyst identifies a wedge

[1] We use an approximation algorithm because finding the guaranteed minimal set of inconsistent constraints requires worst case exponential time.

Figure 2.5: A user interface for specifying shadow constraints: (a) two cast shadows are manually selected on the roof; (b-c) an analyst specifies a point on a cast shadow (red dot) and automatically the range of plausible light directions is rendered (yellow wedge) using the previously entered constraints. Specification of a constraint (c) reduces the width of the wedge to conform to the object casting the shadow (red lines); (d) a constraint may be ignored if the yellow wedge is too narrow to reliably encompass the object which cast the shadow; (e-f) an analyst specifies a point on an attached shadow (red dot) and the orientation and direction of the shadow terminator is automatically suggested by computing the image gradient (pink dotted line and arrow) after which an analyst can accept or adjust the constraint.

that does not overlap the yellow wedge at all, then the new constraint is inconsistent with previous ones and evidences forgery.

An analyst's specification of attached shadow constraints is facilitated by automatically suggesting an attached shadow's orientation at a selected point. Shown in Figure 2.5(e), for example, is a point on an attached shadow (red dot), the automatically rendered terminator orientation (pink line), and direction to the bright side of the shadow (pink arrow). The orientation and direction are determined by simply computing the local intensity gradient in a small spatial neighborhood. The analyst can either accept the suggested attached shadow constraint (Figure 2.5(f)) or rotate it slightly if necessary. Useful attached shadow constraints are found when the half-plane identified by the attached shadow excludes a portion of the yellow region (which does not occur in Figure 2.5(e)).

This simple interface allows an analyst to efficiently identify what is typically a rich set of cast and attached shadow constraints. When cast shadows are well defined, the analyst will typically specify a single wedge, and when cast shadows are ambiguous, the analyst will typically specify multiple wedges that combine to provide a tighter constraint than can be described by a single wedge. When specifying attached shadow constraints, as on clothing, a multitude of redundant constraints may be present, and an expeditious analyst will identify and omit these by using the yellow region as a guide.

2.2 Results

We first validate our technique on a set of large-scale simulations and then provide results from several real-world examples of authentic and visually compelling forgeries.

2.2.1 Simulations

Shown in Figure 2.6 is a computer generated scene with a variety of cast and attached shadows. This scene was rendered multiple times with either an infinitely distant or local point light in one of 49 or 98 locations, respectively. The distant lights were uniformly sampled over a hemisphere excluding an elevation less than 40 degrees. The local lights were uniformly sampled on the same hemisphere and at two different radial distances (R) from the ground plane ($R = 145$ and $R = 265$, where the height of the cylinder in the scene is 30).

We first assess how many wedge and half-plane shadow constraints are necessary to reasonably restrict the space of viable light positions. For each of the 147 computer generated scenes, exemplified by Figure 2.6, a random subset of n cast and attached shadow constraints were first extracted[2]. Then, a random 3-D light position was drawn from a hemisphere,

[2]A subset of n cast and attached shadow constraints were selected from a scene with N constraints as follows. If the total number of subsets of size n was less than 500, then all subsets were considered, otherwise, a random selection of 500 of all possible $N!/n!(N-n)!$ subsets were considered.

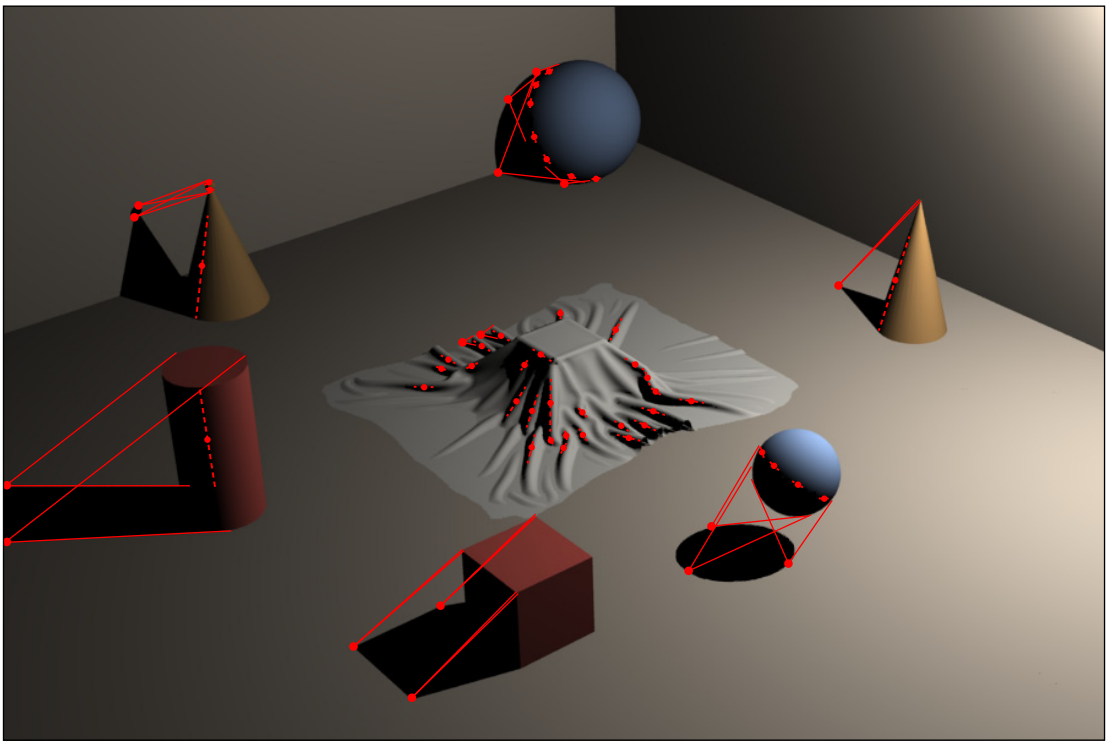


Figure 2.6: A computer generated scene with sixty cast and attached shadows. Wedge constraints from the cast shadows are shown with solid lines. Half-plane constraints from attached shadows are shown with dashed lines.

excluding an elevation less than 10 degrees, and projected into the image plane. Shown in Figure 2.7 (left) is the median probability that this light position was consistent with n constraints (horizontal axis) — the error bars correspond to the 25^{th} and 75^{th} quantile. The difference between the distant and local light is insignificant and so the results are combined.

With only one constraint, the median probability is 0.73 meaning that a single constraint from a falsified part of the image is unlikely to be effective in detecting photo composites. However, this probability decreases quickly as more constraints are added: with three constraints the probability is 0.25, with five constraints the probability is 0.10, with eight constraints the probability is 0.05, and with twenty constraints the probability is 0.01.

We next simulated the creation of a total of 1,008,714 composite photos with inconsistent lighting. Each composite consisted of the seven objects shown in Figure 2.6. A varying sized subset of these objects (94 distinct subsets in total) were illuminated with the light source in one of 147 positions, and the remaining objects were illuminated with a different light source, for a total of 10,731 light pairings. Once composited, a maximum of 1,500 random sets of between 5 and 60 shadow constraints were extracted from each scene to yield approximately 150 million different systems of constraints. These constraints were selected so that the ratio between the wedge cast shadow and the half-plane attached shadow constraints was

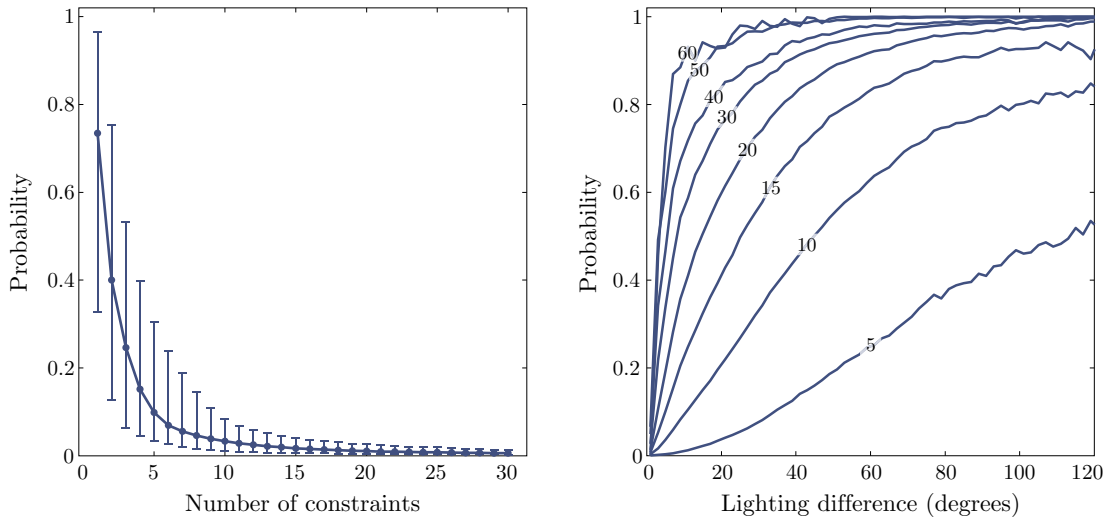


Figure 2.7: Shown on the left is the median probability that a randomly selected light position will satisfy a variable number of cast and attached shadow constraints. The error bars correspond to the 25^{th} and 75^{th} quantile. Shown on the right is the probability of detecting a composite photo as a function of the angular deviation in the position of the light. Each curve corresponds to a different number of shadow constraints.

on average 1:2 (this is the same ratio of constraints found in the original scenes). If no solution is found that satisfies the selected constraints, then the photo is correctly classified as a fake.

Shown in Figure 2.7 (right) is the accuracy with which a forgery was detected in these scenes. The individual curves correspond to the total number of constraints. The horizontal axis corresponds to the angle between the two projected light directions in the composite photo, which is computed as the median angle between the projected directions of the two light sources at each constraint. The detection accuracy improves with the total number of constraints because more constraints carve out an increasingly smaller valid solution space.

With only 10 constraints, the difference in the lighting must reach 100° before the probability of detecting the fake reaches 80%. With 20 constraints, this same accuracy is achieved for a lighting difference of 35°, and with 50 constraints even a small 10° discrepancy can be detected with a probability of 80%. At the same time, the accuracy of correctly classifying an authentic photo in these scenes is 100%. Because any portion of a cast or attached shadow can be used, it is fairly easy to find 50 or more constraints in a typical image.

Shown in Figure 2.8 is an authentic scene with eight constraints. The projected light position is in the intersection of the constraints which is outlined in black. Shown in Figure 2.9 is a falsified version of the scene shown in Figure 2.8 in which the floating sphere from a differently lit scene was inserted. Shown in the top panel are ten constraints selected in this scene and shown in the bottom panel is the result from the automatic detection of a minimal set of unsatisfiable constraints. The red shading corresponds to the positive

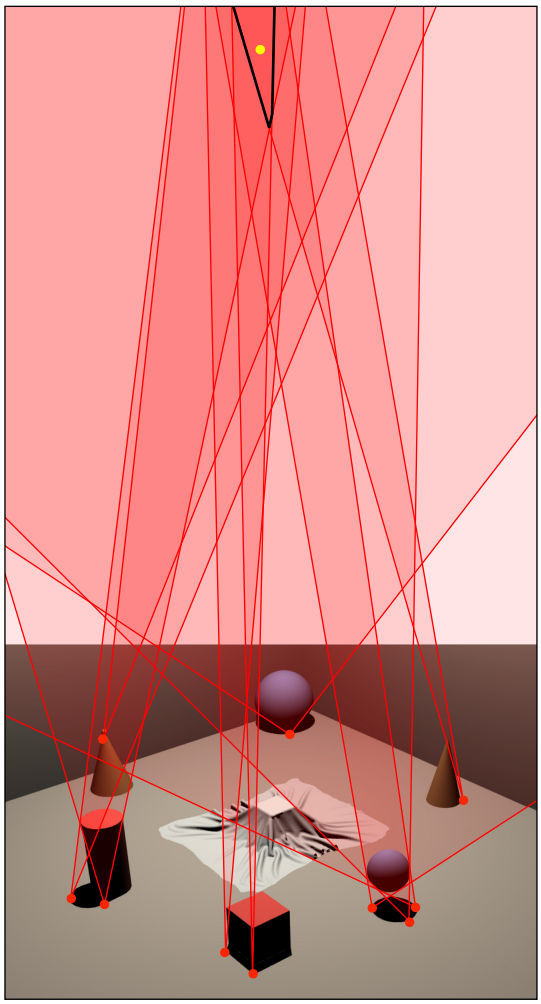


Figure 2.8: Shown is an authentic photo and the successful result of a consistency check. The region of plausible locations for the light source is outlined in black.

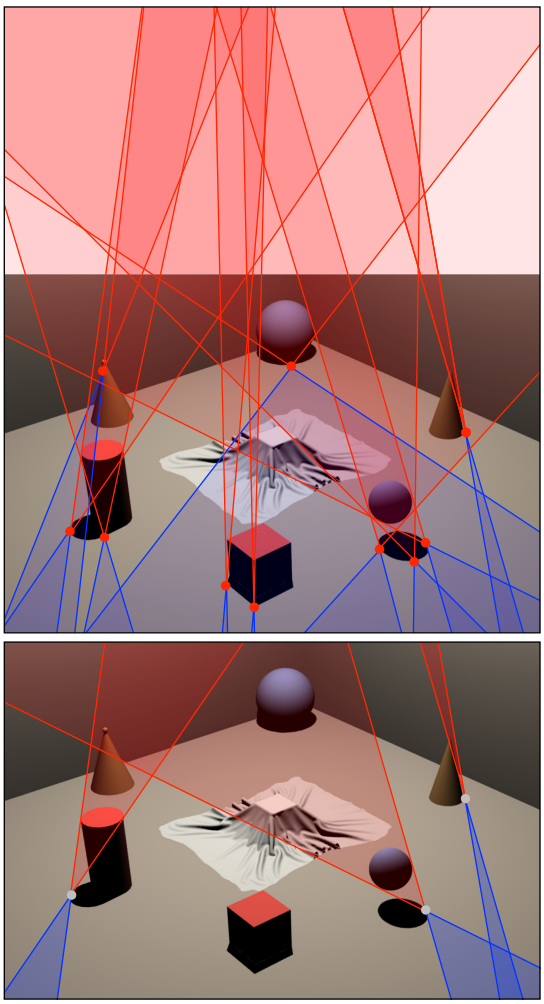


Figure 2.9: Shown is a falsified scene in which the floating sphere's shadow is inconsistent. Top: the complete set of user specified constraints. Bottom: a minimal set of violating constraints is automatically determined from the user specified constraints. Red and blue regions correspond to positive and negative constraints, respectively.

constraints and the blue shading corresponds to the negative constraints — neither are satisfiable. The median angular lighting difference at each constraint is 3.6° which in this case is enough of a difference to create an inconsistency in the shadows.

We further applied our method to several synthetic images from a previously published perceptual study that involved shadows [66]. In that study subjects were presented with relatively simple images in which shadows were cast by two sets of objects onto two separate planar surfaces, and the subjects were asked to indicate if the shadows were mutually consistent (correct) or inconsistent. The results of that study showed that people performed poorly at this task even for these relatively simple scenes containing well-defined shapes with

strong salient features. As shown in Figure 2.10, our method works well on these images that confound human viewers.

2.2.2 Real world

Shown in Figure 2.11 is an authentic photo illuminated with a distant light source (the sun). Also shown is the result of a successful consistency check for eleven cast shadows (dozens of other consistent shadows are available in the scene but their constraints were not rendered to avoid clutter).

Shown in Figure 2.12 is a composite photo illuminated with a local light source where the shadows on the angel are inconsistent with the cloth and the angel's cast shadow. Also shown in this figure is a minimal set of violating constraints. A second composite along with a minimal set of violating constraints is shown in Figure 2.13.

Shown in the top row of Figure 2.14 is a scene illuminated with a local light source. Also shown is the result of a successful consistency check for ten cast and attached shadows (many other consistent shadows are available in the scene but were not rendered to avoid clutter). Shown in the second row of Figure 2.14 is an altered photo in which the duck-billed dinosaur and its shadow were inserted from a separate photo taken with the light in a different position. The panel to the right shows a minimal set of mutually conflicting constraints revealing that the shadow of the duck-billed dinosaur is inconsistent with the rest of the scene. Although significantly different from the correct shadow, the shape and location of the dinosaur's shadow in the composite photo is not, at a casual inspection, obviously fake. Further, subjective arguments about the shape or direction of the shadow would be confounded by the irregular geometry upon which the shadow is cast. In addition to inserting the dinosaur, the image was also modified by rotating the apple and black container counter-clockwise while leaving their shadows unchanged. The apple's shadow, however, remains consistent by our test because the wedge specifying the cast shadow constraint is wide enough to encompass the intersection of the other constraints.

Shown in the bottom row of Figure 2.14 is an altered photo that demonstrates a failure case for our method. The shadow of the robot was manipulated using Photoshop. However the manipulation purposefully shifted the shadow along rays from the light to the original shadow point. The result of this contrived manipulation is a distorted shadow that still generates constraints consistent with the true light location.

Shown in Figure 2.1 is a shadow analysis of a photo from the Apollo 11 moon landing. The constraints correspond to four cast shadows from the astronaut's legs onto the moon's surface, four cast shadows from objects on the moon's surface, and two attached shadows on the astronaut's helmet, and one attached shadow on the footprint on the moon's surface. The region outlined in black identifies locations that satisfy all 11 of these constraints (dozens of other consistent shadows are available in the scene but were not rendered to avoid clutter). Thus we conclude that by our measure the lighting in this image is consistent.

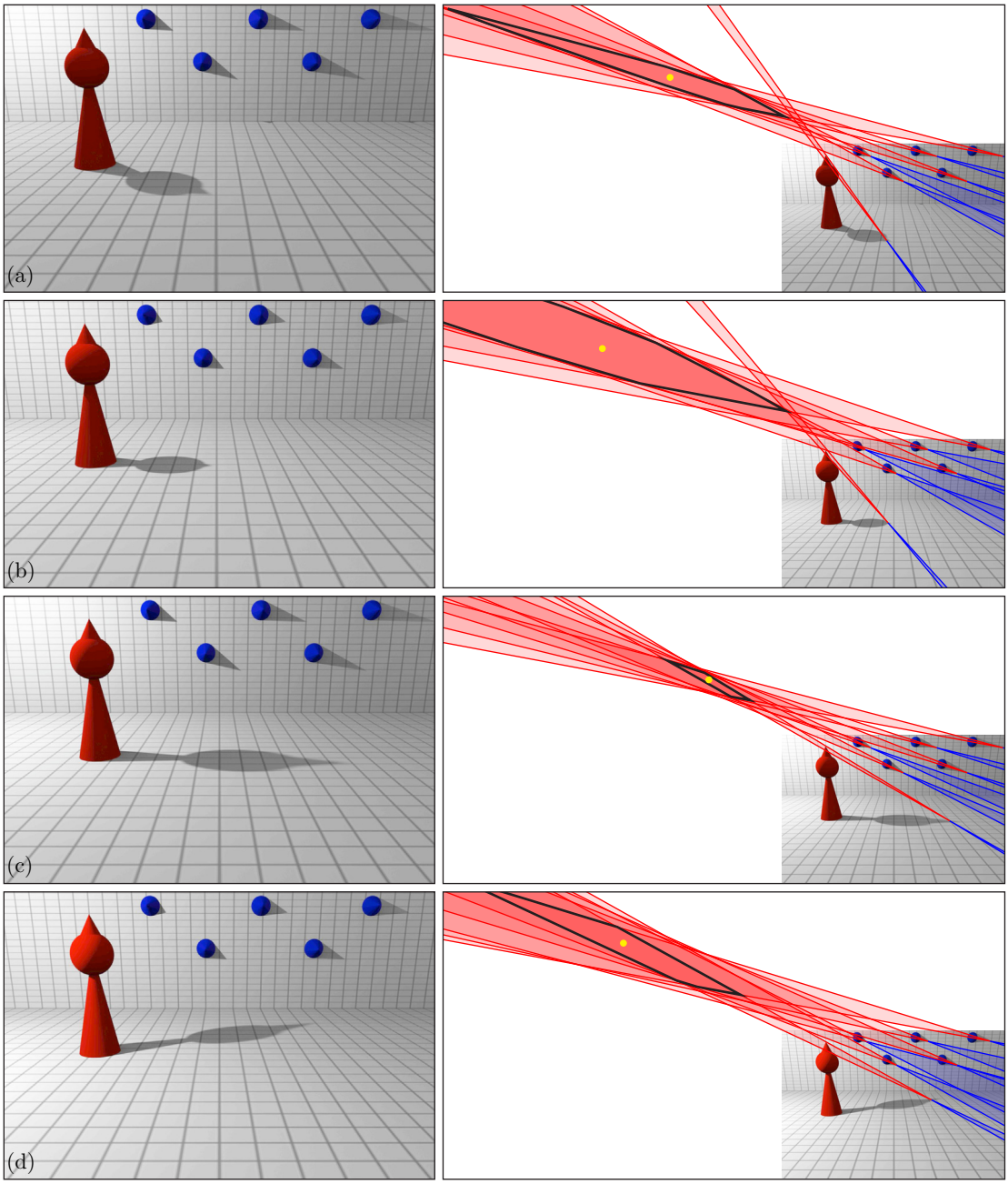
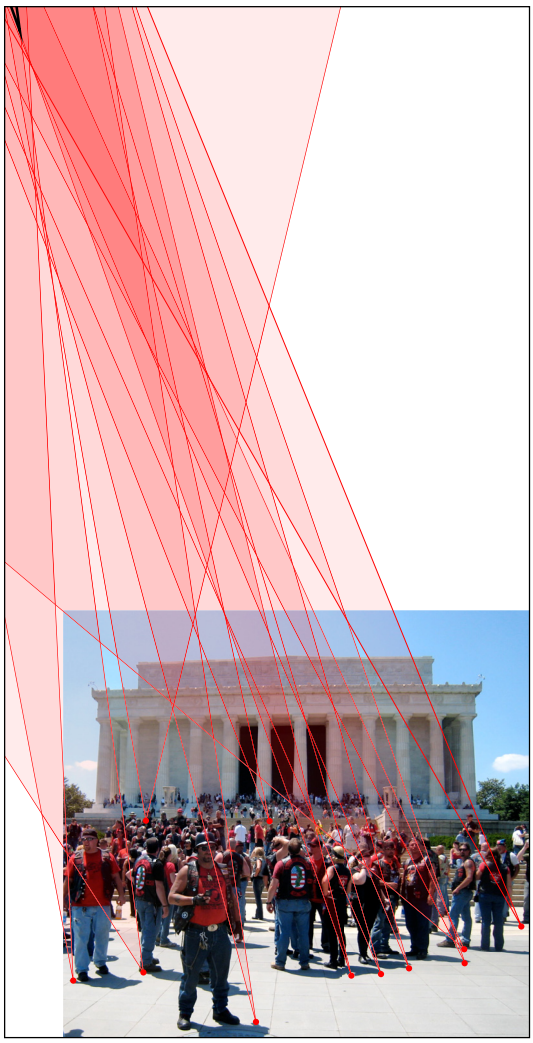
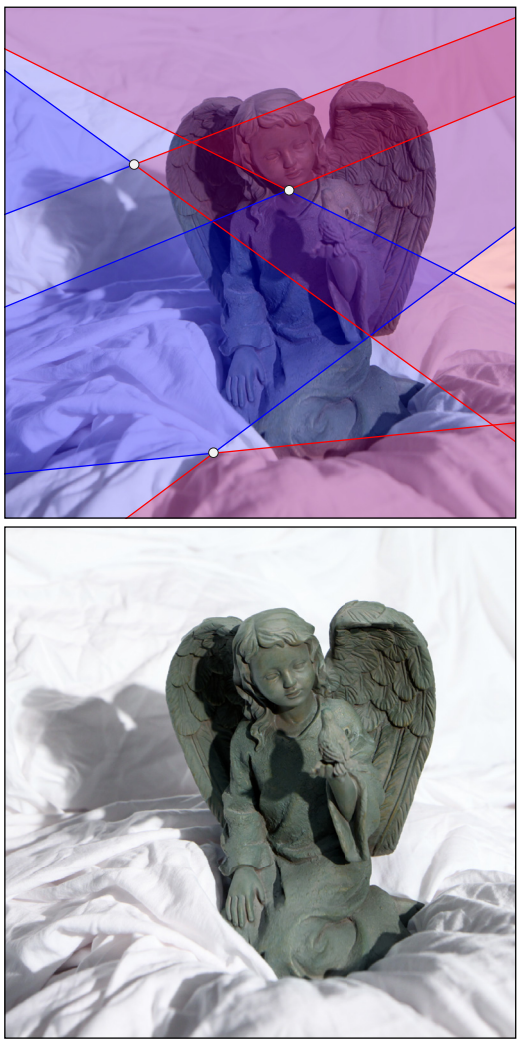


Figure 2.10: Shown in the left column are four sample images used in a perceptual study of observers' ability to judge the consistency of cast shadows [66]. The shadows in panels (a) and (b) are inconsistent with a single light source, while the shadows in panels (c) and (d) are consistent with a single light source. In the original perceptual study, human observers performed at chance level when asked to determine if the shadows in these scenes were consistent or not. Shown in the right column is the result of our analysis in which all four scenes are correctly classified.



Original image copyright 2010, Travlr
http://www.flickr.com/photos/travlr/4654451054

Figure 2.11: An authentic photo and the result of a shadow consistency check. The region of plausible locations for the light source is outlined in black and extends beyond the figure boundary.



Images copyright 2012, Kee, O'Brien, and Farid

Figure 2.12: A forgery and a minimal set of violating constraints (selected from 31 user specified constraints). The shadows cast by the angel's wing, by part of the cloth, and by her chin are mutually inconsistent. Red regions correspond to positive constraints, and blue regions to negative constraints.

2.2.3 User study

We performed a survey of five users with some photo editing experience but no experience with our shadow analysis. Users were trained to specify constraints that tightly encompass the corresponding object, but are conservatively wide enough to defend logically in a post-interview. They were then provided with a training image. In total, training lasted 60 minutes. After training, all five of the users successfully completed the analysis of the image shown in Figure 2.15. The average amount of time taken was 30 minutes. Although

Figure 2.13: A composite photo and a minimal set of violating constraints (these were automatically determined from a total of 25 user specified constraints). The red shaded regions corresponds to the positive constraints and the blue shaded regions correspond to the negative constraints. The three middle panels are a magnified view of the selected shadows.

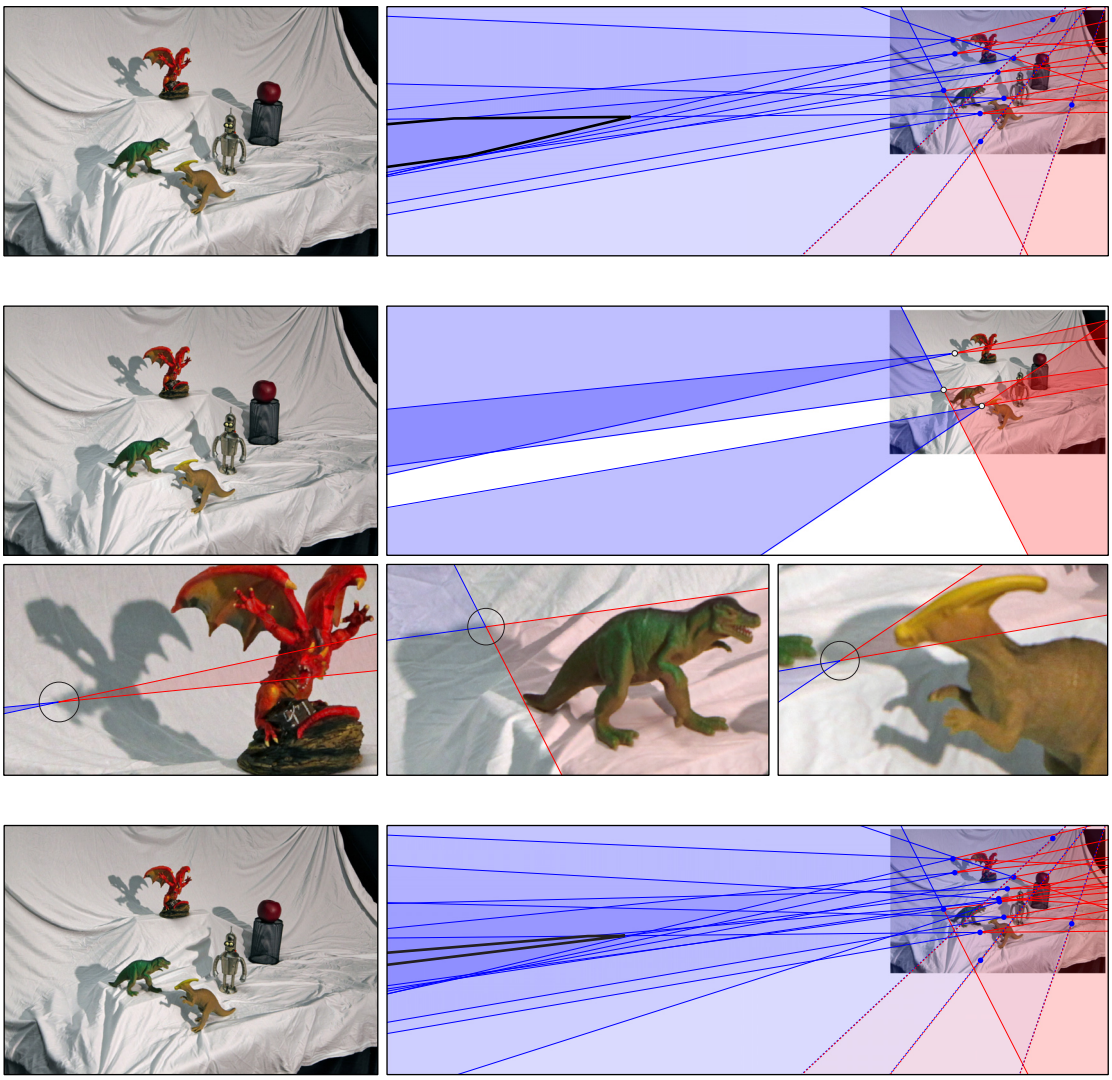


Figure 2.14: An authentic photo (first row) and two fake photos (second and fourth row), and the results of a shadow consistency check. The solid lines correspond to cast shadows and dashed lines correspond to attached shadows. The red shaded regions correspond to the positive constraints and the blue shaded regions correspond to the negative constraints. Where a feasible solution is found, it is outlined in black. The panels in the third row show magnified views of the constraints from the image in the second row. For the authentic image in the top row a valid solution is found. For the fake image in the second row, the right panel shows a minimal set of constraints that are violated revealing that the duck-bill dinosaur's shadow is inconsistent with the rest of the scene (these were automatically determined from a total of 10 user specified constraints). The fake image in the bottom row illustrates a failure case for our method. The robot's shadow has been modified, but the modification still admits a valid solution.

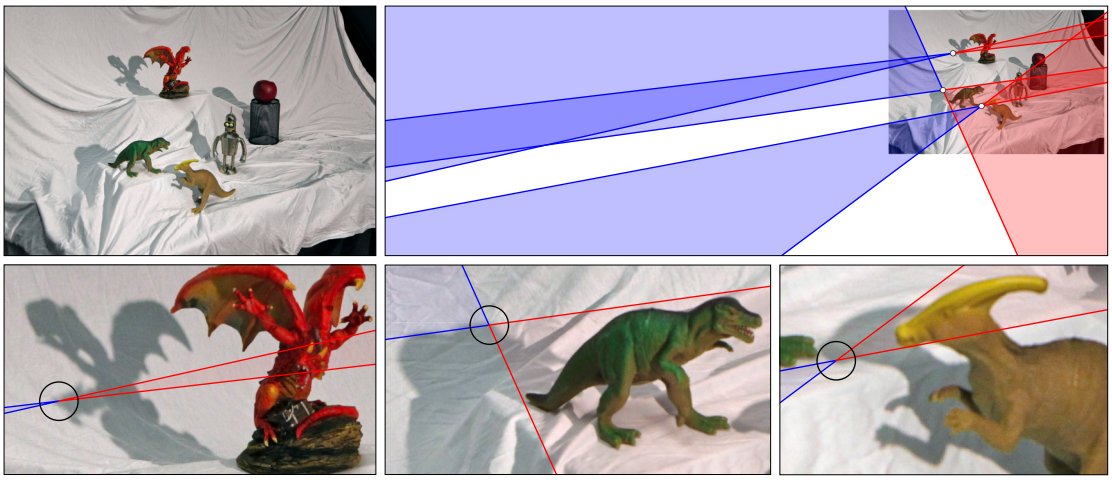


Figure 2.15: Five users were trained to analyze shadows, and were subsequently asked to analyze a forgery (top left). All five users successfully detected the forgery within an average of 30 minutes (top right). Shown are the conflicting shadow constraints input by one of the five users (bottom panels). Note that, although this analysis is similar to that illustrated in Figure 2.14 (center), it was input by a different user.

the fifth user successfully completed the analysis, he employed an invalid constraint. In a post interview, he agreed that it was invalid, and subsequently completed the analysis.

2.2.4 Usability

In certain cases, the specified shadow constraints will restrict the position of the projected light source to be in a closed polygonal region. Although not rendered as such, this is also the case in Figures 2.1, 2.8, and 2.11. In these cases, it is more difficult for a forger to add a consistent shadow. In addition, this restriction to a finite area is a stronger constraint than previous work that identified inconsistencies in the projected 2-D light direction [105].

We note that a local or infinite light source may sometimes project to a point at infinity in the image plane. A local light, for example, will project to infinity if the direction from the camera center to the light is parallel to the image plane. In this case, lines connecting points on a cast shadow to their corresponding points on an object will be parallel and hence intersect at infinity. Shown in Figure 2.16 is such an example where the dashed lines connecting shadow and object on the cone and cube are parallel. An analyst limited by finite image resolution, however, cannot unambiguously identify exact points to specify these parallel lines and instead must specify a broader wedge shape (solid lines). These wedges diverge and therefore their intersection includes finite locations (black region) as well as the true light projection at infinity. Although lights projecting to infinity would create problems for algorithms attempting to find the light position by intersecting nearly parallel rays [234], this situation does not pose any particular difficulty in specifying a valid region using wedge and half-plane constraints.

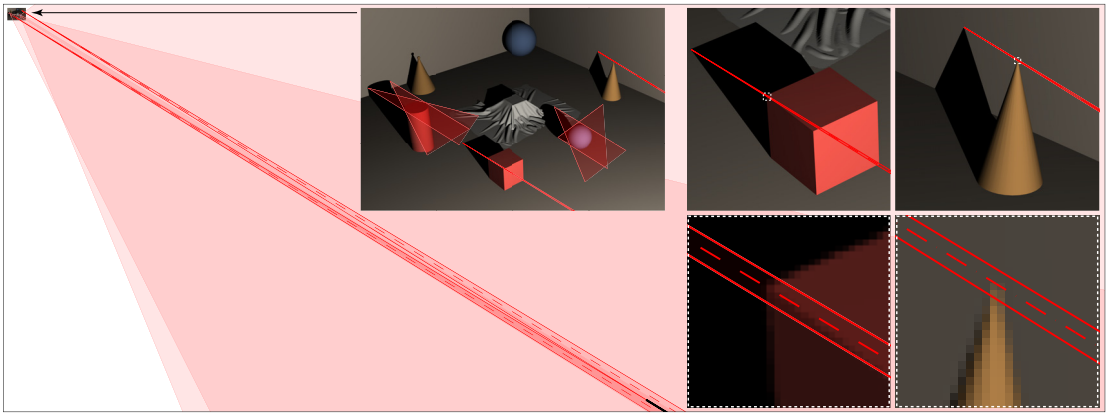


Figure 2.16: A scene is illuminated by a local light source that projects to a point at infinity in the image plane. In this special case, the light casts shadows that project to parallel lines in the image (dashed lines). A forensic analyst, limited by image resolution, must specify broader wedges which diverge (solid red lines), yielding an intersection that includes finite locations (black outline) as well as the true light projection at infinity.

2.2.5 Area lights

Shadows cast by point light sources are common because lighting is often distant from a scene. In many settings, however, light sources occupy a finite area. Such area lights project to a region within an image rather than a point, and preclude the use of cast and attached shadow constraints described in Section 2.1.1 and Section 2.1.2. Similar partial constraints can, however, be used to locate area light sources. We briefly describe these constraints, and present additional results on simulated authentic and falsified photos.

Area lights can be modeled as a summation of individual point lights, Figure 2.17(a). Shown is a synthetic photo in which an area light source has been constructed from a set of small spheres (white dots). Rays from points on this light source (white lines) are occluded by a cylinder, which has been attached to a wall. A distinctive shadow forms, and includes regions of full and partial shadow. Shadowed points that are fully occluded from the light source lie in the shadow's umbra (between the vertical lines); partially shadowed points lie in the penumbra (between the diagonal and vertical lines). The remaining points on the wall are not shadowed by the cylinder (shading varies because the light source is local.)

Wedge-shaped constraints can be specified from within the umbra of a shadow, Figure 2.17(b). Shown in red are two such constraints. Points within the umbra are occluded from every point light source that composes the area light, and wedge constraints from within the umbra therefore identify constraints (red wedges) that are, implicitly, the union of individual point-light constraints from the point in the umbra. The intersection of these umbra-to-object constraints therefore encompasses the area light source. Note, however, that if constraints are placed within the penumbra — rather than the umbra — they will not encompass the light source because penumbra shadows are not occluded from all point lights. (We will return to these penumbra-to-object constraints shortly.)

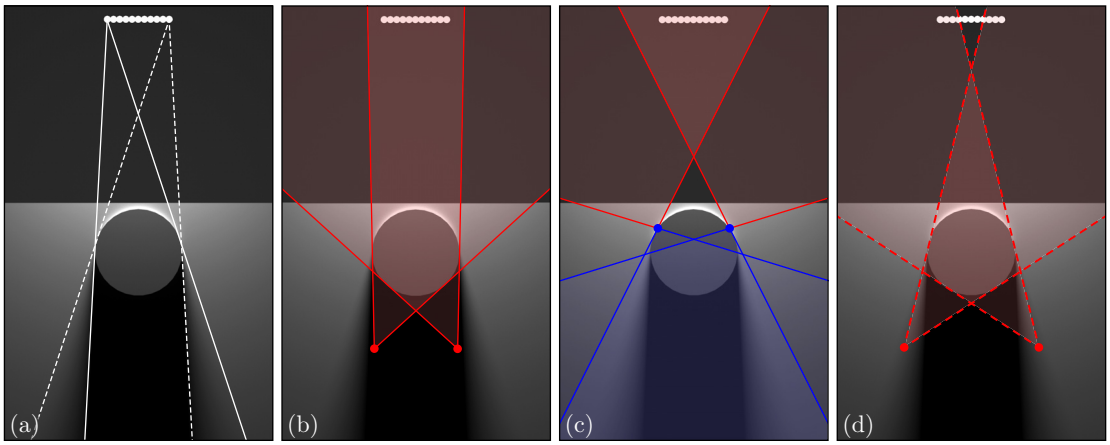


Figure 2.17: Partial constraints from area lights. Shown is a synthetic image in which a cylinder has been attached to a wall and illuminated by an area light source (a). Area lights can be conceptualized as the union of multiple point light sources (white dots), and form distinctive regions of full and partial shadow. Regions of full shadow are called the umbra, and partial shadows are called the penumbra (white lines delineate these regions). Wedge-shaped constraints that contain the area light can be specified from umbra-to-object (b), or from object-to-shadow (c). Their satisfiability can be determined by constructing a linear program. Additional constraints can be specified from penumbra-to-object (d), and must intersect the light source. These constraints verify the authenticity of the penumbra, and their satisfiability can be determined by constructing a separate linear program for each, Section 2.2.5.

Wedge-shaped constraints can also be specified from an object to its shadow, Figure 2.17(c). Shown in blue are two such wedge constraints. Similarly to umbra constraints, these object-to-shadow constraints identify a union of individual point-light constraints. Each point-light constraint identifies points in shadow which may have been occluded by the point on the object. Here the analyst selects a conservatively large wedge constraint that encompasses the entire shadow, including the penumbra. The light source correctly lies within the inverted constraint (red) because it lies in front of the camera.[3]

A variety of cast shadows may be present in scenes illuminated by an area light, and these provide multiple partial constraints on the location of the light. As in the case of point lights, constraints from area lights are combined into a linear program to determine their mutual satisfiability, Section 2.1.3. Unsatisfiable constraints evidence forgery.

An additional and unique constraint can be added when considering scenes that are illuminated by area lights, Figure 2.17(d). Shown are two wedge-shaped constraints that have been specified from points within the penumbra. Such points must be occluded from a portion of the light source. The corresponding constraints must therefore intersect — but will not encompass — the area light (shaded red). Such penubra-to-object constraints can be used to verify that the penumbra is authentic, but must be treated separately from the preceding constraints. Specifically, if the linear program is satisfiable, each additional

[3]When constraints are specified object-to-shadow, lights in front of the camera project into the inverted constraint, Section 2.1.3.

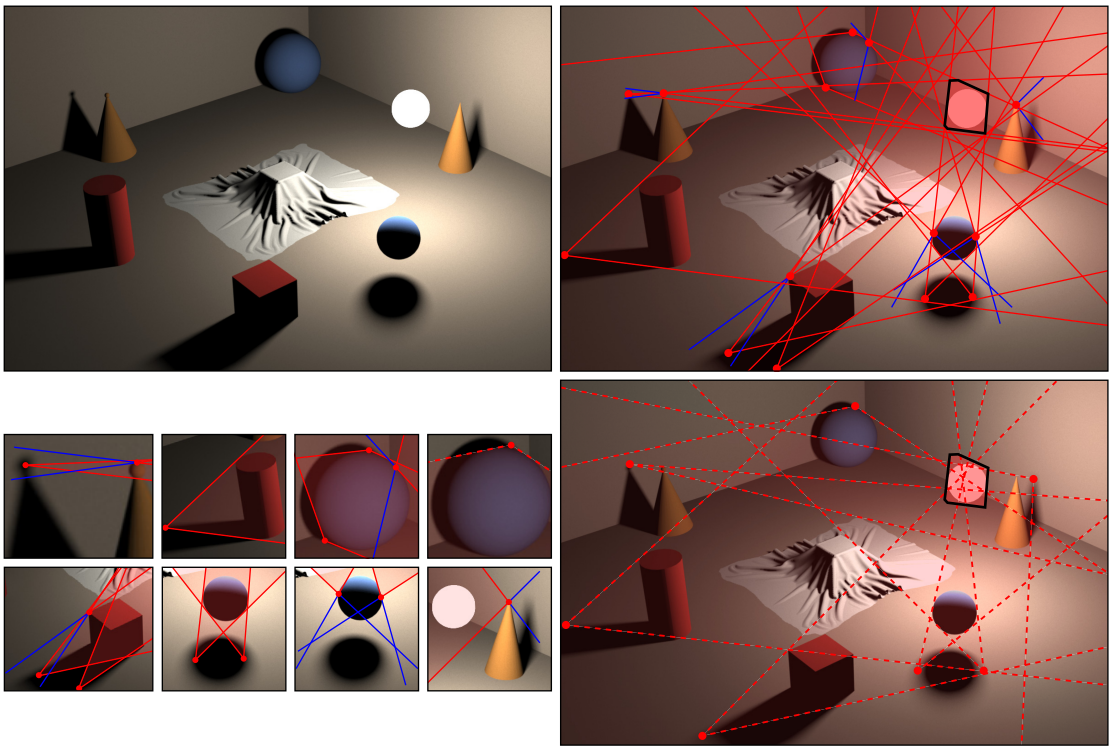


Figure 2.18: An authentic scene is illuminated by a local area light source (top left). A variety of umbra-to-object and object-to-shadow constraints are specified (bottom left). These constraints intersect to identify a feasible region that surrounds the light source (top right, black outline). This partially verifies the consistency of the shadows. Lastly, the authenticity of penumbra regions is tested by specifying additional constraints from penumbra-to-object (bottom right, dashed lines). Each such constraint intersects the feasible region, and the photo is therefore correctly determined to be authentic.

penumbra-to-object constraint is separately added to the linear program and the satisfiability is retested. If any of these additional linear programs are unsatisfiable, forgery is detected. This procedure benefits from the computational efficiency of the linear formulation.

Results

We validate the efficacy of shadow constraints under area lights by simulation. Shown in Figure 2.18 (top left) is an authentic photo of the simulation scene used in Section 2.2.1. A spherical area light (white circle) was placed within the scene, and a variety of umbra-to-object and object-to-shadow constraints were specified (bottom left, solid lines). These constraints are combined (top right), and specify a non-empty feasible region that encompasses the light source (black outline). Additional penumbra-to-object constraints are subsequently specified. For example, one such constraint can be specified on the large sphere (bottom left), and is indicated by dashed lines. A variety of penumbra-to-object constraints were specified (bottom right), and each intersects the feasible region. We therefore correctly determine that the shadows are physically plausible.

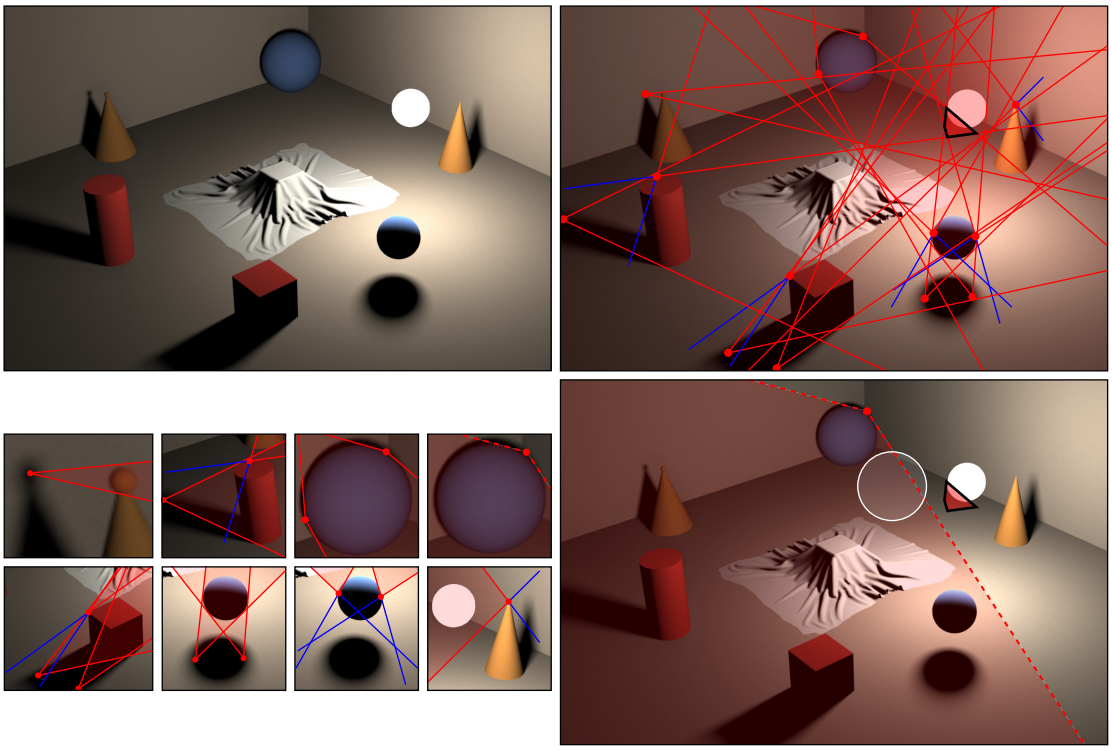


Figure 2.19: A falsified photo is shown in which the shadows cast by the cylinder, left cone, and large sphere have been cast by a different area light source (top left). A variety of umbra-to-object and object-to-shadow constraints are specified (bottom left). These constraints intersect to specify a feasible region that partially includes the known region of the light source (top right, black outline). Additional constraints are tested to determine the authenticity of the penumbra regions. A penumbra-to-object constraint is specified from the penumbra of the large sphere, and this constraint does not intersect the feasible region (bottom right). The photo is therefore determined to be fake. The location of the light source that cast the falsified shadows is also shown (white transparent circle), and demonstrates that the additional penubra-to-object constraint is consistent with its true light source (it intersects the white transparent circle).

Shown in Figure 2.19 (top left) is a falsified photo of the simulation scene. The shadows cast by the cylinder, left-hand cone, and large sphere have been replaced with shadows that were cast by a different area light source. A variety of umbra-to-object and object-to-shadow constraints are specified (bottom left), and are satisfiable (top right) as indicated by the feasible region (black outline). Additional penumbra-to-object constraints are therefore tested. A variety of such constraints are available, and in this case the penumbra-to-object constraint on the large sphere (bottom left, dashed lines) does not intersect the feasible region (bottom right). The shadows in the falsified scene can therefore be correctly determined to be physically implausible under the assumption of a single area light. The area light source that cast the falsified shadows (transparent white outline) is similar in size and location to the authentic light but can nonetheless be detected.

2.3 Discussion

We have described a geometric technique for detecting photo manipulation based on inconsistent shadows. Previous lighting- and shadow-based forensic techniques exploited cast shadows for which a unique object-to-shadow pairing could be estimated (e.g., the tip of a cone). In contrast, partial constraints allow us to exploit a broad range of ambiguous cast and attached shadows, and shadows cast by area lights. Although each partial constraint is typically not highly informative, a combination of many such constraints can be highly effective in detecting inconsistent shadows that are not perceptually obvious. The subsequent determination of shadow consistency is framed as a standard linear programming problem affording a computationally efficient solution. A consistent solution to all specified shadow constraints is interpreted to mean that the shadows are physically plausible while a failure to find a solution is used as evidence of photo tampering. In the latter case, it can be difficult to visually identify which constraints are inconsistent. We therefore developed a method to identify the inconsistent constraints thus providing insight into which parts of an image were manipulated.

This method is intended for use where there is only a single dominant light source. While this limitation does preclude analysis of scenes lit by multiple point lights, it includes the common situation of outdoor scenes lit by the sun or indoor scenes photographed with a flash. Scenes with multiple light sources or strong interreflections are usually evident by a corresponding multitude of shadows for a single object.

Beyond an assumption of linear perspective projection and a single dominant light source, no other assumptions about the scene geometry or photometry are required. Should lens distortion be an issue, standard techniques can be employed to estimate and remove lens distortion. Alternatively the constraint wedges could be expanded to accommodate the bounded movement of features in the image plane due to lens or other types of distortion.

A key step in applying our method is for the analyst to select a set of shadows from the image and to specify appropriate wedges or half-planes. A poor selection of constraints could, of course, lead to a failure in detecting a manipulated image. To minimize the likelihood of this, we have developed a simple user interface which helps an analyst quickly and reliably identify inconsistent shadow constraints (when they are present).

Even with this interface, our method relies on a user correctly selecting constraints. The strength of our approach is that these constraints can be objectively validated because the correspondence between objects and their shadows is generally clear. In cases where the relation is not clear, large encompassing wedges may be specified or other less ambiguous shadows may be used. In the case of area lights, the analyst can select points that lie conservatively within the umbra, and penumbra-to-object constraints can be specified when the area light is sufficiently large for points in the penumbra to be objectively identified. Our method therefore shifts the dialogue from "does the lighting/shadow look correct?" (which is well known to be highly unreliable), to a discussion of whether an analyst has

correctly selected the range of points on an object that correspond to a point on a shadow (a far more objective task). In this regard, our method lets humans do what computers are poor at — understanding scene content — and lets the computer do what humans are poor at — assessing the validity of geometric constraints.

A sufficiently informed forger (as in [221]) could, of course, use this forensic technique to ensure that all shadows are consistent. One way to counter this is to combine this shadow analysis with other techniques for estimating lighting from a single image [185, 187, 135]. This addition will make it more difficult, but never impossible, to create a consistent and visually compelling forgery.

Chapter 3

Shading

We have thus far described how shadows can be used to partially constrain the projected location of point and area light sources. We now describe how the shading on objects provides a similar partial constraint. Specifically, we develop a shading-based partial constraint on the projected location of a point light source in the image plane, and therefore facilitate integration of shading constraints with the shadow constraints described in Chapter 2. These constraints are combined into a simple linear formulation to determine the mutual physical plausibility of shading and shadows in scenes illuminated by point light sources.

Shown in Figure 3.1 is an example of our shading and shadow analysis applied to what might, at first glance, appear to be an unlikely photo. A collection of shading constraints is specified from occluding contours along the objects, and combined with shadow constraints (each shown in red). These constraints partially constrain the projected location of the light source. Because all constraints have a common intersection (outlined in white), we can determine that there exists a point light source that satisfies all constraints. The shading and shadows in Figure 3.1 can therefore be objectively determined to be physically plausible, correctly suggesting that the photo is authentic.

Related Work

A large body of work has been established for the estimation of lighting from an object's shading. Standard techniques often assume, as we do, a single distant point light source illuminating a Lambertian surface of constant reflectance. Within the vision literature, methods include shape-from-shading [268, 93, 122], and inverse lighting in which the 3-D geometry is assumed to be known [174, 266, 91, 269]. Some methods take advantage of both shading and shadow cues, [144, 253]; unlike our approach, these require complete 3-D geometric models. Methods that do not require known 3-D geometry have used manually-approximated 3-D geometry [114] but are sensitive to user input. When 3-D models are not available or cannot be easily estimated, 2-D occluding contours can be used to estimate a 2-D subset of the 3-D light direction [185].

In the forensics literature, shading and shadows have been individually exploited to

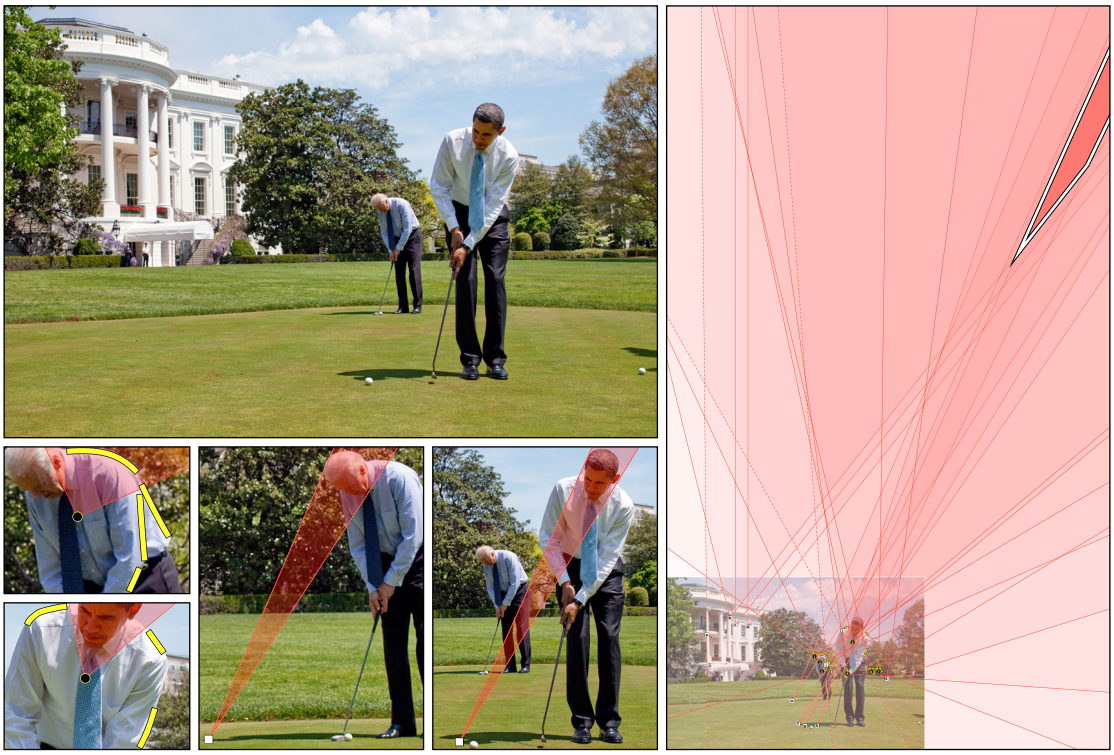


Figure 3.1: Our analysis finds that the shading and shadows in this photo of President Obama and Vice President Biden are physically consistent with a single light source. Shading along occluding contours and cast and attached shadows provide constraints (red wedges) on the projected location of the light source. Shown are four representative examples (magnify right panel for more details). All constraints intersect (white outline), indicating that they are mutually consistent.

detect photo tampering [105, 107, 265, 117, 151, 211]. The combination of shading and shadow has been considered for estimating 2-D light direction [110], and for estimating the location of light sources [234, 235]. This work is however demonstrated only in the analysis of paintings, and assumes orthographic projection (an overly limiting assumption in many practical situations).

Within this context, our work takes a different approach to lighting-based forensics. Rather than explicitly estimating lighting, we use shading and shadows to partially constrain the projected location of the light source, Figure 3.1. This approach has several advantages in that it makes no assumptions about the underlying scene geometry, makes relatively few assumptions about the scene (single distant point light source illuminating a Lambertian surface of constant reflectance), assumes a more plausible model of perspective projection, and combines both shading and shadow cues within a single unified framework. Perhaps most importantly, our method can be applied to a broader class of photos for which the explicit estimation of lighting would be difficult or impossible.

3.1 Methods

In this section we formulate how shading on an object constrains the location of a distant point light source in 3-D and in the projected 2-D image. We show that these 2-D shading constraints can be combined with 2-D shadow constraints within a linear framework to determine if the shading and shadows in an image are physically consistent with a single light source.

Throughout, the following notational conventions are used. Scalar values are lowercase letters (a), vectors are bold lowercase letters ($\mathbf{v}$), and matrices are bold capital letters ($\mathbf{A}$). Italicized vectors denote quantities in the 2-D image plane ($\boldsymbol{v}$), while roman vectors denote quantities in the 3-D world or higher dimensions ($\mathbf{v}$). Individual elements of vectors or columns of matrices are referenced by subscripts: v_k is the k^{th} element of vector $\mathbf{v}$, and $\mathbf{a}_k$ is k^{th} column of matrix $\mathbf{A}$. Vectors that represent homogeneous quantities in the image plane are denoted as $\tilde{\boldsymbol{v}}$ and their Cartesian equivalent is simply $\boldsymbol{v}$.

3.1.1 Shading constraints (3-D)

Under an infinitely distant point light source, the radiance r of an image patch that corresponds to a surface with constant Lambertian reflectance and normal $\mathbf{n} = \begin{pmatrix}\mathrm{n}_x & \mathrm{n}_y & \mathrm{n}_z\end{pmatrix}^T$ is given by

$$r = \mathbf{n} \cdot \mathbf{s} + a, \tag{3.1}$$

where $\mathbf{s} = \begin{pmatrix}\mathrm{s}_x & \mathrm{s}_y & \mathrm{s}_z\end{pmatrix}^T$ specifies the direction to the light source, and the magnitude of $\mathbf{s}$ is proportional to the light's brightness. The constant term a is the so-called ambient term which approximates indirect illumination. Note that this expression assumes that the angle between the surface normal and light is less than 90°. The components of this lighting model can be estimated from $k \geq 4$ patches with known surface normals. The equations for each surface normal and corresponding radiance are packed into the following linear system:

$$\begin{pmatrix} \mathbf{n}_1^T & 1 \\ \mathbf{n}_2^T & 1 \\ \vdots & \vdots \\ \mathbf{n}_k^T & 1 \end{pmatrix} \begin{pmatrix} \mathbf{s} \\ a \end{pmatrix} = \mathbf{r} \tag{3.2}$$

$$\mathbf{N}\mathbf{e} = \mathbf{r}\ , \tag{3.3}$$

where $\mathbf{r}$ is a vector of observed radiance for each patch. The lighting parameters $\mathbf{e}$ can be estimated using standard least squares:

$$\mathbf{e}_0 = (\mathbf{N}^T\,\mathbf{N})^{-1}\mathbf{N}^T\,\mathbf{r}\ , \tag{3.4}$$

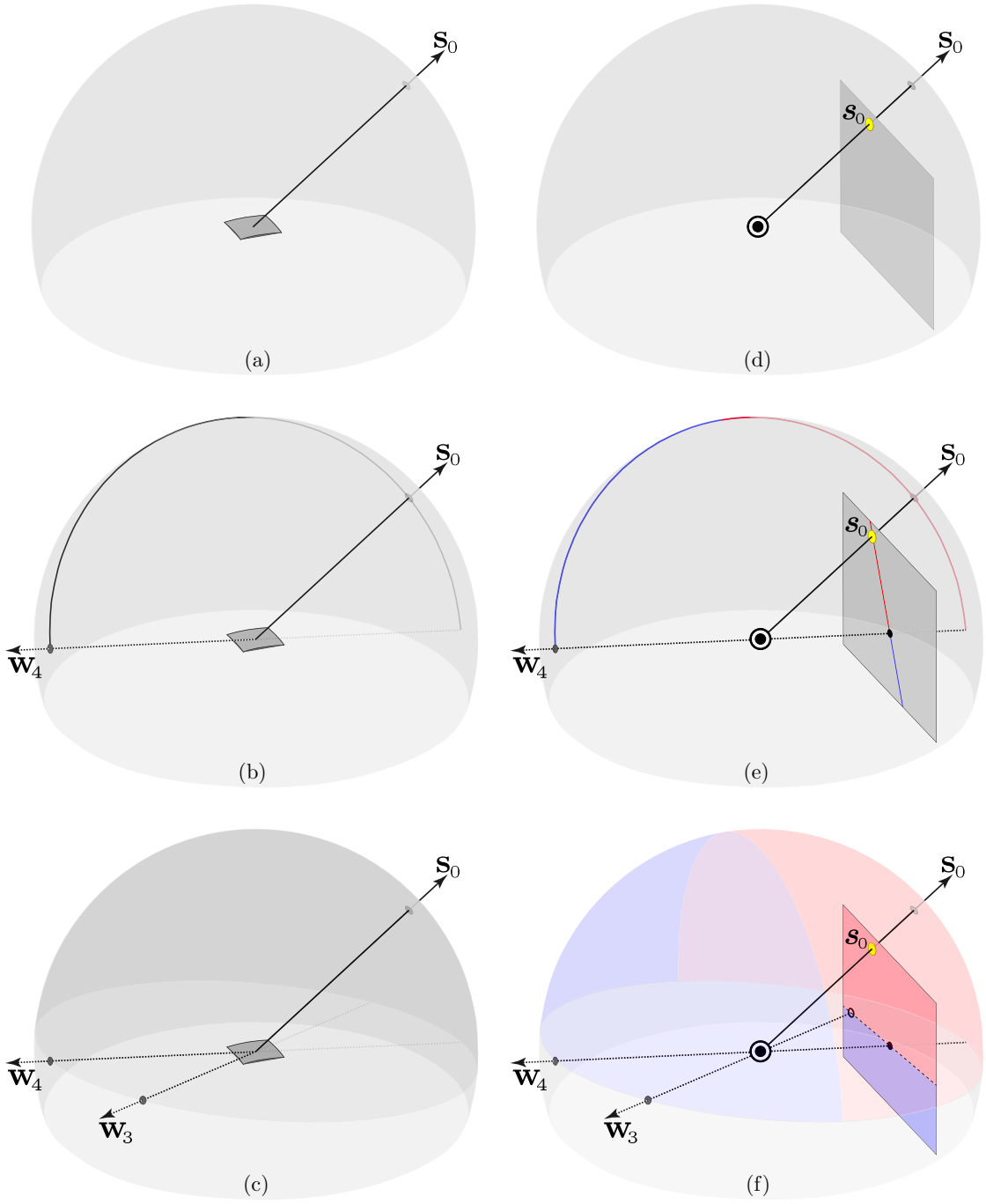


Figure 3.2: Shading constraints in 3-D (left) and 2-D (right): (a) when the linear system in Equation (3.3) is fully conditioned, the 3-D location of the light source, $\mathbf{s}_0$, can be uniquely determined; (b-c) if the linear system is not fully conditioned, then the 3-D location of the light source is constrained to a semi-circular arc or a hemisphere, depending on the rank of the linear system ($\mathbf{w}_3$ and $\mathbf{w}_4$ are the vectors corresponding to the singular values of the linear system). Shown in panels (d)-(f) is the projection of these constraints onto the image plane, where the camera's center of projection is denoted by ◉.

where $\mathbf{e}_0$ denotes the estimated parameters, and its first three components specify the brightness-scaled light direction, $\mathbf{s}_0$. The normalized light direction, $\mathbf{s}_0/||\mathbf{s}_0||$, can be visualized as a point on the surface of the unit sphere, Figure 3.2(a).

When the matrix $\mathbf{N}$ is well conditioned, this estimation is straightforward. In practice, however, the matrix $\mathbf{N}$ may be poorly conditioned or rank deficient due to a limited number, or range, of known surface normals. In such situations it is not possible to determine an exact solution, $\mathbf{e}_0$, of the lighting parameters but it is still possible to constrain the solution to a space of possible solutions.

Consider the singular value decomposition (SVD) of the linear system in Equation (3.3):

$$\mathbf{U}\mathbf{\Lambda}\mathbf{V}^T\,\mathbf{e} \;=\; \mathbf{r} \quad . \tag{3.5}$$

This decomposition projects the lighting parameters $\mathbf{e}$ onto an orthonormal basis $\mathbf{V}$:

$$\mathbf{V} = \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{v}_4 \end{pmatrix} = \begin{pmatrix} \mathbf{w}_1 & \mathbf{w}_2 & \mathbf{w}_3 & \mathbf{w}_4 \\ b_1 & b_2 & b_3 & b_4 \end{pmatrix} \tag{3.6}$$

where $\mathbf{w}_i$ are 3×1 vectors and b_i are scalar values. Note that $\mathbf{w}_i$ and b_i form a basis for the brightness-scaled light direction $\mathbf{s}$ and ambient component a. After projection onto vectors $\mathbf{v}_i$, $\mathbf{e}$ is scaled by the diagonal matrix $\mathbf{\Lambda}$ of singular values λ_i, and then transformed into image radiance by the matrix $\mathbf{U}$. Using the convention that the singular values, $\lambda_1, \ldots, \lambda_4$, are sorted from largest to smallest, the well-conditioned dimensions of the solution space are given by the columns of $\mathbf{\Lambda}$ and $\mathbf{V}$ that correspond to singular values $\lambda_i/\lambda_1 \gg 0$. The lighting parameters $\mathbf{e}$ can be estimated within the subspace spanned by the well-conditioned vectors $\mathbf{v}_i$ of the singular value decomposition:

$$\mathbf{e}_0 = \mathbf{N}^{+}\mathbf{r} \quad , \tag{3.7}$$

where $\mathbf{N}^{+}$ denotes the pseudo-inverse.

When one singular value is degenerate, $\lambda_4/\lambda_1 \approx 0$, possible lighting solutions are given by $\mathbf{e}_0 + \alpha\mathbf{v}_4$, where $\mathbf{e}_0$ is the solution to the modified linear system in Equation (3.7), α is a scalar, and $\mathbf{v}_4$ is the fourth column of matrix $\mathbf{V}$. This constrains the brightness-scaled direction of the light source $\mathbf{s}$ to a line in 3-D:

$$\mathbf{s} = \mathbf{s}_0 + \alpha\mathbf{w}_4 \quad , \tag{3.8}$$

where $\mathbf{s}_0$ is the first three components of $\mathbf{e}_0$ corresponding to the light direction, and $\mathbf{w}_4$ is the first three components of singular vector $\mathbf{v}_4$. This line passes through the point $\mathbf{s}_0$ and is parallel to $\mathbf{w}_4$. The corresponding set of normalized light directions, $\mathbf{s}/||\mathbf{s}||$, forms a semi-circular arc, which is the projection of the line onto the unit sphere, Figure 3.2(b).

When two singular values are degenerate, $\lambda_4/\lambda_1 \approx 0$ and $\lambda_3/\lambda_1 \approx 0$, possible lighting

solutions are given by $\mathbf{e}_0 + \alpha\mathbf{v}_4 + \beta\mathbf{v}_3$, where $\mathbf{e}_0$ is the least-squares solution to the modified linear system, and both α and β are scalars. This constrains the brightness-scaled direction of the light source $\mathbf{s}$ to a plane in 3-D:

$$\mathbf{s} = \mathbf{s}_0 + \alpha\mathbf{w}_4 + \beta\mathbf{w}_3 \quad , \tag{3.9}$$

which passes through $\mathbf{s}_0$ and is parallel to both $\mathbf{w}_4$ and $\mathbf{w}_3$. The corresponding set of normalized light directions forms a hemispherical surface, which is the projection of the plane onto the unit sphere, Figure 3.2(c).

When three singular values are degenerate, $\lambda_4/\lambda_1 \approx 0$, $\lambda_3/\lambda_1 \approx 0$, and $\lambda_2/\lambda_1 \approx 0$, the null space of the linear system spans the full 3-D space, thus providing no constraint on the light direction $\mathbf{s}$.

3.1.2 Shading constraints (2-D)

In the previous section we described how an object's shading may constrain the 3-D direction of a distant point light. In this section we develop a representation of these constraints that can be combined with constraints from shadows (detailed in Section 3.1.4). Light directions in 3-D can be treated as points at infinity and projected into the 2-D image plane. Constraints on the 3-D lighting direction can, therefore, be expressed as constraints on the projected light location in the 2-D image plane.

For simplicity, but without loss of generality, let the surface normals $\mathbf{n}$ and direction of the light source $\mathbf{s}$ be specified in the camera's coordinate system in which the camera faces down the $-z$ axis. Under an idealized perspective camera with focal length f and image center $\boldsymbol{c}$, the light source at infinity in direction $\mathbf{s}$ projects to a point in the image plane:

$$\tilde{\boldsymbol{s}} = \begin{pmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & -1 \end{pmatrix} \mathbf{s} \tag{3.10}$$

$$= \mathbf{K}\mathbf{s} \tag{3.11}$$

where the image point $\tilde{\boldsymbol{s}}$ is specified in homogeneous coordinates, and $\mathbf{K}$ is the camera's intrinsic matrix.

When the linear system, Equation (3.3), is fully conditioned, the direction of the light source, $\mathbf{s}$, can be unambiguously estimated, $\mathbf{s} = \mathbf{s}_0$, Figure 3.2(a). This constrains the projected location, $\tilde{\boldsymbol{s}}$, of the light source $\mathbf{s}$ to a single point in the image:

$$\tilde{\boldsymbol{s}} = \mathbf{K}\mathbf{s}_0 \quad . \tag{3.12}$$

Shown in Figure 3.2(d) is the projection, $\boldsymbol{s}_0$, of the 3-D direction, $\mathbf{s}_0$, onto a 2-D image plane (the camera center of projection is denoted by ◉).

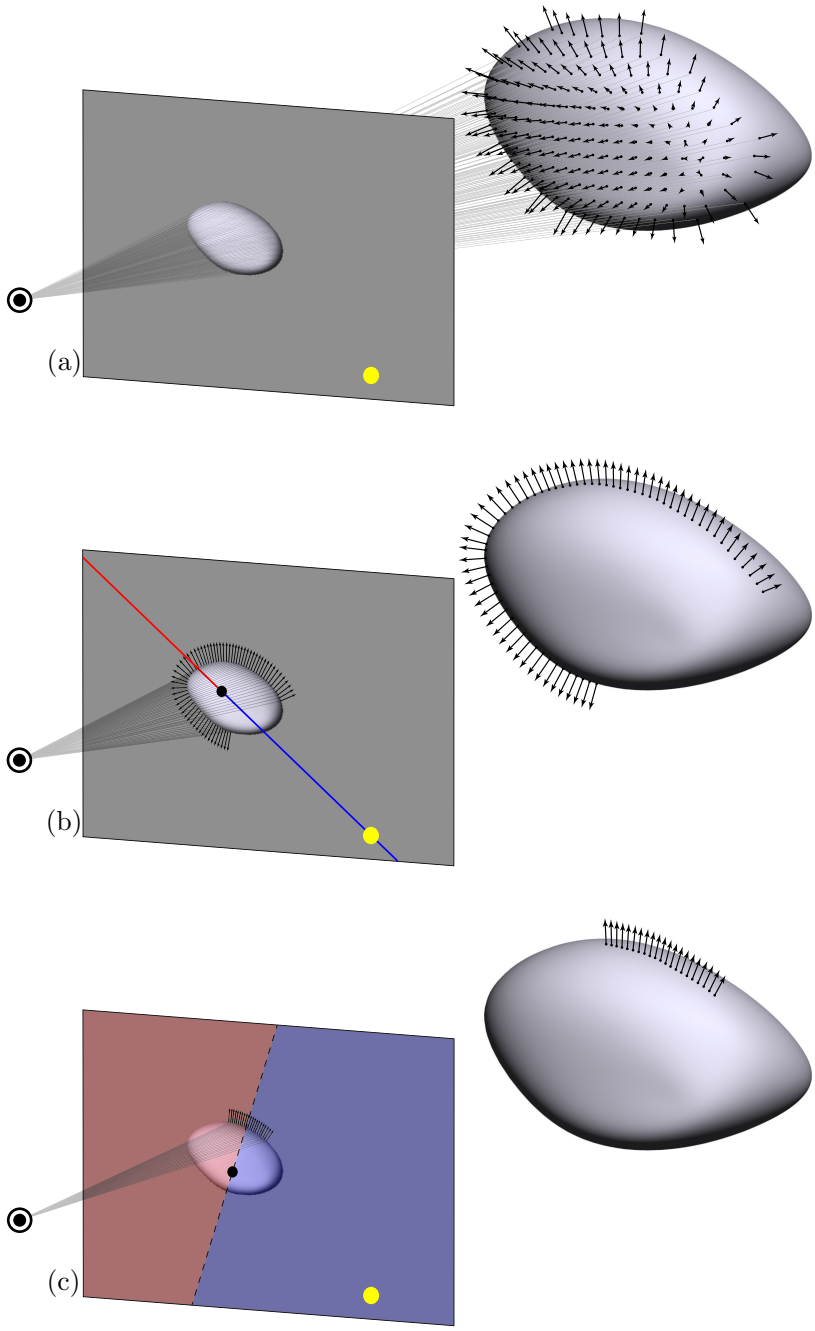


Figure 3.3: Geometric configurations that lead to point (a), line (b), and half-plane (c) constraints on the projected location of the light source. Shown in each panel is the projection of a 3-D object (right) into the image plane (the camera's center of projection is denoted by ◉). A point constraint can be specified when a full 3-D model is available. Surface normals along an extended portion of an occluding contour provide a line constraint, and normals along a small portion of the occluding contour provide a half-plane constraint.

When one singular value is degenerate, the brightness-scaled light direction is constrained to a line in 3-D, $\mathbf{s} = \mathbf{s}_0 + \alpha\mathbf{w}_4$, which corresponds to a half-circle of directions, Figure 3.2(b). These directions project to a line in the image plane:

$$\tilde{\boldsymbol{s}} \;=\; \mathbf{K}(\mathbf{s}_0 + \alpha\mathbf{w}_4) \quad . \tag{3.13}$$

Shown in Figure 3.2(e) is the projection of the 3-D constraint (a half-circle of directions) onto a 2-D image plane. Note that light directions behind the camera's center of projection (blue arc) are projected into the lower part of the line (blue line segment), while light directions in front of the camera (red arc) are projected into the upper part of the line (red line segment). These two line segments meet at the image of vector $\mathbf{w}_4$, $\mathbf{K}\mathbf{w}_4$ (denoted by $\bullet$ in the image plane). This line and its segments can be described by a position $\boldsymbol{p}$ and direction $\boldsymbol{d}$ in the image:

$$\tilde{\boldsymbol{p}} \;=\; \mathbf{K}\mathbf{w}_4 \tag{3.14}$$

$$\boldsymbol{d} \;=\; \mathrm{sign}\,(\mathrm{s}_{0_z})(\boldsymbol{p} - \boldsymbol{s}_0) \quad , \tag{3.15}$$

where $\boldsymbol{s}_0$ is the projected location of the least-squares solution of the light direction, $\mathbf{s}_0$, and $\mathrm{s}_{0_z}$is its z-component. Direction $\boldsymbol{d}$ identifies the line segment onto which light directions in front of the camera project (see Appendix A.1 for a derivation).

When two singular values are degenerate, the brightness-scaled light direction is constrained to a plane in 3-D, $\mathbf{s} = \mathbf{s}_0 + \alpha\mathbf{w}_4 + \beta\mathbf{w}_3$, which corresponds to a half-sphere of directions, Figure 3.2(c). These directions project to a plane in the image:

$$\tilde{\boldsymbol{s}} \;=\; \mathbf{K}(\mathbf{s}_0 + \alpha\mathbf{w}_4 + \beta\mathbf{w}_3) \quad . \tag{3.16}$$

Although this plane encompasses the entire image, it consists of two half-planes, Figure 3.2(f): one for light directions in front of the camera (red), and one for directions behind the camera (blue). These half-planes are described by a position $\boldsymbol{p}$ and direction $\boldsymbol{d}$ in the image:

$$\tilde{\boldsymbol{p}} \;=\; \mathbf{K}\mathbf{w}_4 \tag{3.17}$$

$$\boldsymbol{d} \;=\; \mathrm{sign}\left(\tilde{\boldsymbol{h}} \cdot \tilde{\boldsymbol{s}}_0\right)\begin{pmatrix}\tilde{h}_1 & \tilde{h}_2\end{pmatrix}^T \quad , \tag{3.18}$$

where $\tilde{\boldsymbol{h}}$ is the homogeneous line through the image of $\mathbf{w}_4$ and $\mathbf{w}_3$:

$$\tilde{\boldsymbol{h}} \;=\; \mathbf{K}\mathbf{w}_4 \times \mathbf{K}\mathbf{w}_3 \quad . \tag{3.19}$$

Direction $\boldsymbol{d}$ identifies the half-plane into which light directions in front of the camera project (see Appendix A.1 for a derivation).

Lastly, when three singular values are degenerate, the light direction is unconstrained in 3-D and thus provides no constraint on the projected location of the light in the image

plane.

Figure 3.3 illustrates how different configurations of surface normals can lead to the point, line, and half-plane constraints on the projected location of the light source. Shown in panel (a) is the case where a full 3-D model of the object is known, yielding a full rank matrix $\mathbf{N}$ in Equation (3.3) and a point constraint. Shown in panel (b) is the case where only approximately coplanar normals along an extended portion of the object's occluding boundary are known. These occluding normals yield one degenerate singular value in $\mathbf{N}$, producing a line constraint. Lastly, shown in panel (c) is a case where only normals along a small portion of the object's occluding boundary are specified, yielding two degenerate singular values in $\mathbf{N}$ and a half-plane constraint.

3.1.3 Satisfiability of multiple constraints

In a forensic setting, we would like to determine if the shading on objects in a photo are physically consistent with a single light source. We therefore combine the shading constraints from multiple objects in a scene into a single linear system to determine if they are collectively satisfiable. Consider first the case in which the light source is in front of the camera. When the light position can be constrained to a point in the image, Equation (3.12), the following equality constraint (in homogeneous coordinates) is placed on the location of the light source $\boldsymbol{s}$:

$$\tilde{\boldsymbol{s}} \;=\; \mathbf{K}\mathbf{s}_0 \quad . \tag{3.20}$$

When the light position can be constrained to a line, Equation (3.13), the following constraint is placed on $\boldsymbol{s}$:

$$\boldsymbol{s} \;=\; \alpha\boldsymbol{d} + \boldsymbol{p} \tag{3.21}$$

$$\alpha \;\geq\; 0 \quad , \tag{3.22}$$

where $\boldsymbol{p}$ and $\boldsymbol{d}$ are given by Equations (3.14)-(3.15). And, when the light position is constrained by a half-plane, Equation (3.16), an inequality constraint is placed on $\boldsymbol{s}$:

$$\boldsymbol{s}\cdot\boldsymbol{d} - \boldsymbol{p}\cdot\boldsymbol{d} \;\geq\; 0 \quad , \tag{3.23}$$

where $\boldsymbol{p}$ and $\boldsymbol{d}$ are given by Equations (3.17)-(3.19).

Each of these constraints are linear, and multiple constraints can therefore be combined into a linear program to check for mutual consistency. A viable solution to the linear program means that the shading in an image is consistent, while a failure to find a viable solution means that one or more of the constraints are inconsistent with a single light source.

When the light source is behind the camera, its projection into the image is inverted by linear perspective projection. To contend with this possibility, a second linear program is

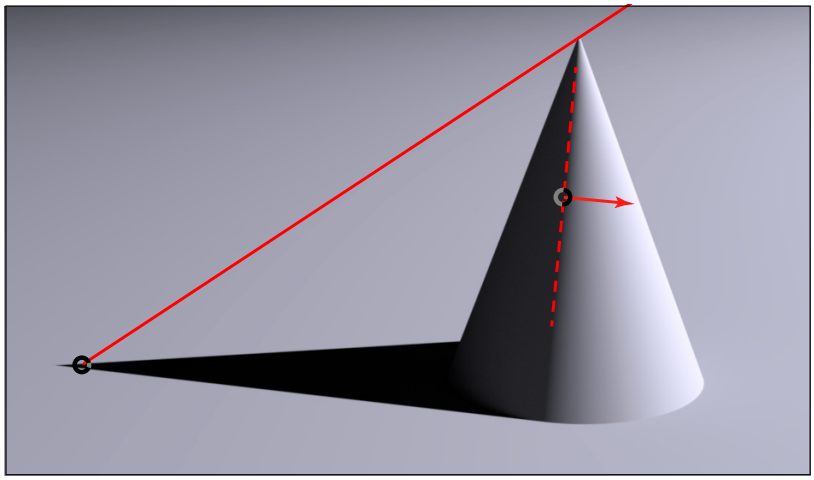


Figure 3.4: A cast and attached shadow constraint. The projected location of the light source lies on a line that connects a point in the cone's cast shadow with its corresponding point on the cone. An attached shadow on the cone constrains the projected location of the light source to be in a half-plane, specified by the orientation and polarity of the shadow terminator.

constructed in which all constraint inequalities are reversed. The shading is determined to be consistent if either linear program is satisfiable.

3.1.4 Shadow constraints (2-D)

The previous sections have described how the shading of an object can be used to constrain the projected location of a point light source. As described in Chapter 2, cast and attached shadows also constrain the projected location of a point light source. These 2-D shadow constraints take the form of linear inequalities and can therefore be combined with our shading constraints to further constrain the projected location of the light source. We briefly restate the constraints provided by cast and attached shadows to reveal their relationship with shading constraints.

Cast shadows

Fundamentally, cast shadow constraints are defined by connecting a point in shadow to its corresponding point on the shadow-casting object, as illustrated in Figure 3.4 (solid line). In the 3-D scene, this ray intersects the light source. In a 2-D image of the scene created under linear perspective, the projection of this ray remains a straight line that must connect the images of the shadow and object points, and must intersect the projected location of the light source. These constraints hold regardless of the geometry of the object and the surface onto which the shadows are cast, and hold for either an infinitely distant or local light. The theoretical partial constraint on the light source, $\boldsymbol{s}$, that cast shadows specify is

therefore a line:

$$\boldsymbol{s} = \alpha \boldsymbol{d} + \boldsymbol{p} \tag{3.24}$$

$$\alpha \geq 0 \quad , \tag{3.25}$$

where $\boldsymbol{p}$ is a point in a shadow, and $\boldsymbol{d}$ is a vector connecting $\boldsymbol{p}$ to the corresponding point on the object. Notice that this constraint has the same form as the line constraint provided by shading, Equations (3.21)-(3.22). As such, these cast shadow constraints can be seamlessly combined with the shading constraints, Section 3.1.3.

Attached shadows

Attached shadows occur when objects occlude the light from themselves, leaving a portion of the object in shadow, as illustrated in Figure 3.4 (dashed line and arrow). Points are in shadow if the surface normal makes an angle greater than 90 degrees with the direction toward the light. The surface contour that divides points in shadow from those receiving light is known as the terminator. Surface normals along this contour form a 90 degree angle with the direction to the light. Similar to a cast shadow, there is a correspondence between points in and out of shadow on either side of the terminator. This correspondence, however, can only be specified to within a half-plane because the light's elevation is ambiguous to within 180 degrees. Attached shadows therefore constrain the projected location of the light source $\boldsymbol{s}$ to a half-plane:

$$\boldsymbol{s} \cdot \boldsymbol{d} - \boldsymbol{p} \cdot \boldsymbol{d} \geq 0 \quad , \tag{3.26}$$

where $\boldsymbol{p}$ is a point on a terminator, and $\boldsymbol{d}$ is a vector orthogonal to the terminator in the direction of the illuminated side of the attached shadow. Notice that this constraint has the same form as the half-plane constraint provided by shading, Equation (3.23). As such, these attached shadow constraints can be seamlessly combined with the shading constraints, Section 3.1.3.

As with the shading constraints, when the light source is behind the camera, its projection into the image is inverted by linear perspective projection. To contend with this possibility, a second linear program is constructed in which all constraint inequalities are reversed. The shadows are determined to be consistent if either linear program is satisfiable.

3.1.5 Uncertain shading constraints

To this point we have assumed an ideal system in which the shading and shadow constraints can be precisely and accurately specified. However, luminance non-linearities, imprecise normals due to limited pixel resolution, and deviations from the Lambertian and distant point-source assumptions will introduce errors into the estimated light direction. It is

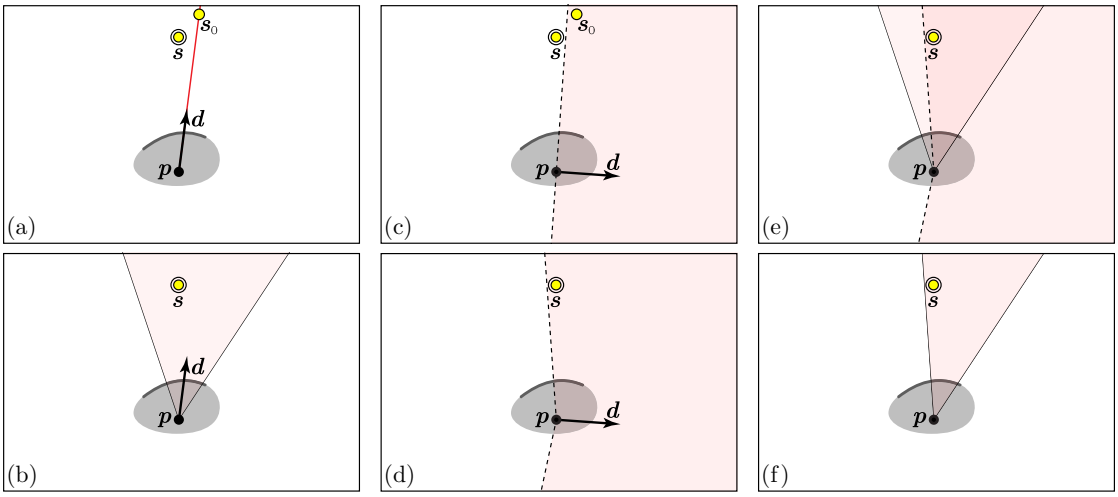


Figure 3.5: Shown in each panel is an occluding contour along the top of an object. The projected location of the light source is denoted as **s** and the least squares estimate of the light source is denoted as $\mathbf{s}_0$. Shown are the following constraints, parameterzied by a position **p** and direction **d**: (a) a line constraint; (b) its corresponding convex wedge-shaped constraint; (c) a half-plane constraint; (d) its corresponding concave wedge-shaped constraint; (e) the pair of wedge-shaped constraints superimposed; and (f) the intersection of these two constraints yielding a single convex constraint.

therefore important to incorporate a model of uncertainty into the estimation of lighting. We note that in a forensic setting, a full 3-D model of an object is not typically available. As such, we will focus only on the more commonly occurring situation in which the lighting can only be constrained to a line or half-plane.

For simplicity, and because it yields linear constraints, we assume that uncertainty in the shading constraints affects only the constraint direction $\boldsymbol{d}$ in Equations (3.21) and (3.23) and that $\boldsymbol{p}$ is not significantly affected. This corresponds to the assumption that the dominant source of uncertainty is in the measured image radiance.

Shown in Figure 3.5(a) and (c) are examples where the estimated lighting constraints (parameterized by $\boldsymbol{p}$ and $\boldsymbol{d}$) do not correctly constrain the true projected location of the light source, **s**. By introducing uncertainty in the estimated direction $\boldsymbol{d}$, these constraints take on a wedge-shape as shown in Figure 3.5(b) and (d). In each case, the wedge-shaped region specifies a range of plausible projected light positions, and contends with uncertainty in the initial estimated lighting constraints.

Convex wedges, Figure 3.5(b), can be described by the intersection of two half-planes, each of which are linear. Concave wedges, Figure 3.5(d), however, are not linear, precluding a linear solution to determine the satisfiability of constraints, Section 3.1.3.

This problem of concave constraints can be overcome by noting that, although the line and half-plane constraints are described in terms of degeneracy in the linear system, both constraints can be constructed for a particular linear system. That is, instead of constructing only one constraint for each object (either a line or half-plane constraint), we construct both constraints regardless of conditioning, and add angular uncertainty as dictated by the

corresponding singular values of the linear system, Figure 3.5(e). (More details on this construction are given below). The constraints are then intersected to produce a combined constraint, Figure 3.5(f). If this combined constraint specifies a convex wedge, it can be incorporated into the linear program, Section 3.1.3, otherwise, it can conservatively omitted.

The uncertainty of constraints, characterized as the width of the resulting wedge-shaped region, specify a confidence on the angle between the constraint direction $\boldsymbol{d}$ and the vector from $\boldsymbol{p}$ to the light source (up to an inversion of $\boldsymbol{d}$ due to lights that lie behind the camera). In part, this uncertainty depends upon measurable characteristics of the linear system, which includes the singular value associated with the constraint, the range of surface normals that are available, and the extent to which the estimated lighting parameters, $\mathbf{e}$, can explain the observed image intensity.

We therefore describe the uncertainty for line and half-plane constraints as functions $\Phi_{\text{L}}(\cdot)$ and $\Phi_{\text{H}}(\cdot)$, of three parameters κ, θ, and ρ. The first parameter, κ, is the condition of the line constraint, $\kappa_{\text{L}} = \lambda_3/\lambda_1$, or half-plane constraint, $\kappa_{\text{H}} = \lambda_2/\lambda_1$. The second parameter, θ, is the range of surface normals, measured to be independent of image resolution by constructing a histogram over the directions of the normals and computing the number of degrees spanned by non-empty bins. Lastly, the extent to which the lighting model can explain the observed image intensity is measured as the R-value, ρ, of the least-squares solution.

Large-scale simulations are used to estimate the uncertainty functions, and are detailed in Appendix A.2. Random 3-D shapes are imaged under random point lights, and by cameras with random focal lengths and image centers. Common sources of uncertainty are included: a secondary point light is added to simulate inter-reflections and other effects, and the images are subjected to noise, random gamma corrections, and quantization. Appendix A.3 details this shading. Random portions of the shaded objects are selected to compute the directions $\boldsymbol{d}$ of the line and half-plane constraints, and their angles ϕ to the light source. This simulation is repeated for many such samples ϕ, and each is assigned to its requisite quantized location in the domain of the uncertainty functions. The uncertainty functions specify the angle below which a majority of the angles ϕ fall at each quantized location in the function domain (e.g. 99.9%).

To build a constraint from an occluding contour, the position, direction, and width of the line and half-plane wedges are estimated, and their intersection computed. This produces a combined wedge with half-width, Φ_{C}, direction $\boldsymbol{d}_c$, and position $\boldsymbol{p}_c = \boldsymbol{p}$. Convex wedges ($\Phi_{\text{C}} \leq 90°$) can be described by a pair of half-planes:

$$\boldsymbol{s} \cdot \boldsymbol{n}_{\text{C}}^1 - \boldsymbol{p}_{\text{C}} \cdot \boldsymbol{n}_{\text{C}}^1 \geq 0 \quad \text{and} \quad \boldsymbol{s} \cdot \boldsymbol{n}_{\text{C}}^2 - \boldsymbol{p}_{\text{C}} \cdot \boldsymbol{n}_{\text{C}}^2 \geq 0 \quad , \tag{3.27}$$

where $\boldsymbol{s}$ is the unknown location of the light source. Vectors $\boldsymbol{n}_{\text{C}}^i$ are

$$\boldsymbol{n}_{\text{C}}^1 = \mathbf{H}^T \, \mathbf{R} \boldsymbol{d}_{\text{C}} \quad \text{and} \quad \boldsymbol{n}_{\text{C}}^2 = \mathbf{H}\mathbf{R}^T \, \boldsymbol{d}_{\text{C}} \quad , \tag{3.28}$$

where $\mathbf{H}$ is a 90° rotation matrix, and matrix $\mathbf{R}$ specifies a rotation of Φ_C degrees. Lastly, concave wedges ($\Phi_C > 90°$) cannot be described by linear constraints, and are therefore omitted for simplicity.

3.1.6 Uncertain shadow constraints

It may be difficult to accurately specify a cast shadow constraint due to finite resolution, ambiguity in matching a cast shadow to its corresponding object, or when the shadow is indistinct. As described in Chapter 2, modeling uncertainty in cast shadow constraints is relatively straightforward. We briefly restate this model here for completeness.

The line constraint for cast shadows can be relaxed to a wedge-shaped constraint by allowing a forensic analyst to select a point on a shadow and a range of possible corresponding points on the object that safely encompass the correct object-shadow pairing. These relaxed constraints can be described by a pair of half-planes:

$$\boldsymbol{s} \cdot \boldsymbol{n}_i^1 - \boldsymbol{p}_i \cdot \boldsymbol{n}_i^1 \geq 0 \quad \text{and} \quad \boldsymbol{s} \cdot \boldsymbol{n}_i^2 - \boldsymbol{p}_i \cdot \boldsymbol{n}_i^2 \geq 0 \quad , \tag{3.29}$$

where the normal vectors $\boldsymbol{n}_i$ are defined such that points within the wedge are on the positive side of both half-plane constraints. These wedge constraints are linear and the satisfiability of a collection of these, along with other linear shading-based constraints, can be determined using linear programming, Section 3.1.3. Note that these constraints are identical to the wedge-shaped shading constraints, Equation (3.27). Lastly, an analyst may sometimes prefer to instead select a point on an object and a range of possible corresponding points in shadow, rather than a point in shadow and a range of possible points on the object. In this case the vectors $\boldsymbol{n}$ in Equation (3.29) are simply replaced by $-\boldsymbol{n}$.

As noted previously, we assume that an attached shadow constraint can be reliably specified due to the relative ease with which the terminator on an attached shadow can be specified. This has the benefit of avoiding non-linear concave constraints which would preclude a linear solution, Section 3.1.3.

3.1.7 Forensics

A photo may contain a variety of shaded objects along with cast and attached shadows. Shadow constraints are specified manually, while shading constraints are specified semi-automatically. In practice full 3-D geometric models of scene objects are rarely available and determining surface normals from a single image is generally difficult and error prone. We therefore restrict ourselves to using points along occluding contours, for which normals can be reasonably estimated. The normals along occluding contours will generally not span the full 3-D space and therefore produce under-constrained systems corresponding to line and half-plane shading constraints.

An analyst specifies a shading constraint by first identifying a segment of an occluding

contour having a constant Lambertian reflectance. The 2-D image normals, $\boldsymbol{n} = \begin{pmatrix} n_x & n_y \end{pmatrix}^T$, are computed from this occluding contour (a Bezier or other general-purpose curve can be fit to this contour, from which the 2-D normals can be analytically estimated). Under a model of linear perspective projection, the corresponding 3-D surface normal is given by:

$$\mathbf{n} = \begin{bmatrix} n_x \\ n_y \\ \frac{1}{f}\,\boldsymbol{n} \cdot (\boldsymbol{x} - \boldsymbol{c}) \end{bmatrix}, \tag{3.30}$$

where $\boldsymbol{x}$ is the 2-D image location of each corresponding 2-D normal, f is the camera's focal length, $\boldsymbol{c}$ is the image center, and it is assumed that $\mathbf{n}$ is subsequently scaled to unit length (see Appendix A.4 for a derivation of this expression). These normals and the corresponding image radiance r along the occluding contour are used to estimate the lighting parameters $\mathbf{e}$, Equation (3.2). If the camera's focal length, f, and image center, $\boldsymbol{c}$, are known or can be estimated, then this estimation is straightforward. As we will describe below, however, explicit knowledge of these parameters is not necessary.

The linear system used to estimate the lighting parameters $\mathbf{e}$, Equation (3.2), will typically be degenerate due to the limited range of normals provided by an occluding contour. This system, therefore, is solved by computing the pseudo-inverse, Equation (3.7). The SVD of the matrix of surface normals, $\mathbf{N} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T$, is then computed. Using the estimated lighting parameters, $\mathbf{e}_0 = (\mathbf{s}_0 \;\; a_0)^T$, and the matrix $\mathbf{V}$, the position and direction of the line constraint, $\boldsymbol{p}_{\text{L}}$ and $\boldsymbol{d}_{\text{L}}$, are given by Equations (3.14)-(3.15), and the position and direction of the half-plane constraint, $\boldsymbol{p}_{\text{H}}$ and $\boldsymbol{d}_{\text{H}}$, are given by Equations (3.17)-(3.18).

Next, an occluding contour's line and half-plane constraints are adjusted to contend with uncertainty in the estimation of the lighting parameters, Section 3.1.5. The line constraint takes on a wedge shape with position $\boldsymbol{p}_{\text{L}}$, direction $\boldsymbol{d}_{\text{L}}$, and angular half-width $\Phi_{\text{L}}(\kappa_{\text{L}}, \theta, \rho)$; the half-plane constraint also takes on a wedge shape with position $\boldsymbol{p}_{\text{H}}$, direction $\boldsymbol{d}_{\text{H}}$, and angular half-width $\Phi_{\text{H}}(\kappa_{\text{H}}, \theta, \rho)$. These two constraints are intersected to produce a single wedge-shaped region with position $\boldsymbol{p}_{\text{C}} = \boldsymbol{p}_{\text{L}} = \boldsymbol{p}_{\text{H}}$. Because both constituent wedges may be concave, the direction $\boldsymbol{d}_{\text{C}}$, and angular half-width, Φ_{C}, of the combined wedge is computed using a simple case-based algorithm. If the resulting wedge is concave, and hence non-linear, no constraint is specified.

Each convex constraint $(\boldsymbol{p}_c,\ \boldsymbol{d}_c,\ \Phi_{\text{C}})$ is added to a linear program by constructing a pair of half-planes, Equations (3.27)-(3.28). Similarly, each shadow constraint, specified by the analyst as either a wedge or half-plane constraint is added to the linear program. A solution to the linear program means that the shading and shadows in an image are consistent, while a failure to find a viable solution means that one or more of the constraints are inconsistent with a single light source. In order to contend with the situation when the light source is

behind the camera, each constraint is reversed to yield a second linear program. Both of these systems are solved to determine if there is a viable light location that satisfies all of the constraints.

Lastly, recall that the estimation of the 3-D surface normals, Equation (3.30), requires knowledge of the camera's focal length f and image center $\boldsymbol{c}$. Because the estimation of these parameters from a single image can be difficult or, in some situations, impossible, we forgo their explicit estimation and instead simply search for camera parameters that yield a viable solution to the specified constraints. Specifically, we quantize the range of plausible focal lengths and image centers into a small set of values. The functions $\Phi_{\text{L}}(\cdot)$ and $\Phi_{\text{H}}(\cdot)$, which account for the uncertainty in the estimation of the lighting parameters, are adjusted to include the additional uncertainty that is introduced by this quantization. Each quantized tuple of camera parameters is used to construct the required 3-D surface normals, and the satisfiability of the corresponding linear programs is checked. This approach is computationally efficient due to the simplicity of the linear formulation, and because a relatively coarse quantization of the focal length and image center has proven to be sufficient, Appendix A.2.

3.2 Results

We partition our results into two main sections. In the first section we report on a set of large-scale simulations used to validate the model of uncertainty employed by the shading-based constraints, Section 3.1.5. In particular, we report on the reliability and specificity of the shading constraints under a variety of assumptions that are consistent with, and that violate, our basic imaging assumptions. In the second section we apply our forensic technique to authentic and visually compelling forgeries.

3.2.1 Simulations

In the ideal situation, the wedge-shaped shading constraints will be just broad enough so as to always encompass the estimated light position, but narrow enough so as to reliably detect differences in lighting. We quantify these criteria in terms of the reliability and specificity of the wedge-shaped constraints. The reliability is measured as the probability that a constraint correctly encompasses the true light source. The specificity is measured as the probability that a random light position satisfies one or more constraints.

Reliability

Shown in Figure 3.6 is a scene rendered with the physically-based renderer Mitsuba [104]. This rock garden scene was rendered with either a point light, an area light subtending 5°, or a physically-based daylight model [208]. In each case, one of these distant light sources was placed in one of 49 possible directions, uniformly sampled over a hemisphere (excluding

Figure 3.6: A synthetic scene rendered with the physically based renderer Mitsuba [104]. This basic scene was used to test the reliability and specificity of the shading constraints.

an elevation below 40 degrees). This sampling was chosen to match that used in simulation of shadow constraints, Chapter 2.

The scene was also rendered with the objects having different material properties. Each object's material was of constant reflectance and was either Lambertian, a rough diffuse material (modeling surfaces such as concrete and clay [188]), a rough plastic material (modeled by the physically-based Beckmann macrofacet distribution [14]), or a woven cloth material [102].

The image rendering was configured to yield objects with an average resolution of 276 (default) or 136 square pixels, 1% (default) or 2% additive noise, and gamma correction in the range $[1.2,\ 1.8]$ (default) or $[1.5,\ 2.1]$.

Shading constraints were constructed by dividing the illuminated portion of each rock's occluding contour into 20 segments, from which a random subset of between 1 and 20 segments were selected. The corresponding 3-D occluding normals were estimated by selecting a focal length and image center within precision assumed by the uncertainty functions, Φ_{L} and Φ_{H}, Appendix A.2. A total of $784,000$ constraints were specified for each of 9 types of scenes (3 lighting scenarios, 3 material types, and 3 rendering configurations).

The constraints were designed to have an average reliability of 99.8%, Appendix A.2. We therefore expect that this percentage of constraints will encompass the projected location of the light source.

In the case of a distant point light source, Lambertian reflectance, a 276 square pixel resolution, 1% additive noise, and gamma in the range [1.2, 1.8], the projected location of the light source was encompassed by 99.3% of the constraints. An area light yielded the same reliability, and a daylight model yielded a higher reliability of 99.9%. The daylight model improves reliability because it contributes a stronger ambient illumination that decreases the signal to noise ratio. This, in turn, results in wider constraints thus improving overall reliability.

In the case of non-Lambertian surfaces illuminated by a distant point light source and the default remaining imaging parameters, rough diffuse objects, rough plastic, and cloth material yielded constraints with a reliability of 99.6%, 96.4%, and 99.2% respectively. Rough plastic led to a reduction in reliability by several percentage points because the roughened specular component biases the estimated light direction, but does not yield a corresponding increase in the constraint width.

Returning to a distant point light source and a Lambertian surface, a decrease in resolution to 136 square pixels yielded a reliability of 94.4%. An increase in the additive noise to 2% yielded a reliability of 99.8%. An increase in gamma into the range [1.5, 2.1] yielded a reliability of 99.4%. As expected, the reduced resolution impacts the reliability. The increased gamma had little affect on reliability, and interestingly an increased level of noise improves reliability. This higher reliability is due to a reduced signal to noise ratio, which reduces the R-value of the least-squares solution and therefore yields a wider constraint.

In summary, the overall reliability of the shading constraints was computed by averaging constraints across all scene types (lights, reflectance models, and image degradations). This yields an overall reliability of 99.3%.

Specificity

The width of each constraint determines its specificity, which is defined as the likelihood that a random light direction is satisfied by one, or more, constraints. The specificity is computed for a variable number and range of occluding contours from the scenes described in the preceding section, Figure 3.6.

Random 3-D light directions were drawn from a hemisphere, excluding an elevation below 10 degrees and projected into the image.[1] When the entire illuminated portion of an occluding contour was used, the median constraint width was 30°, and was satisfied by 19% of random light directions. As constraints were added, up to a total of 8, this probability decreased to 6%, Figure 3.7.

When one half of the contour was selected, the median constraint width was 75°, and was satisfied by 44% of random light directions. This probability decreased to 18% when all 8 available constraints were used, Figure 3.7. When one quarter of the contour was selected, the median constraint width was 119°, and was satisfied by 65% of random light directions.

[1] This sampling affords comparison with the specificity of simulated shadow constraints, Figure 2.7 (left).

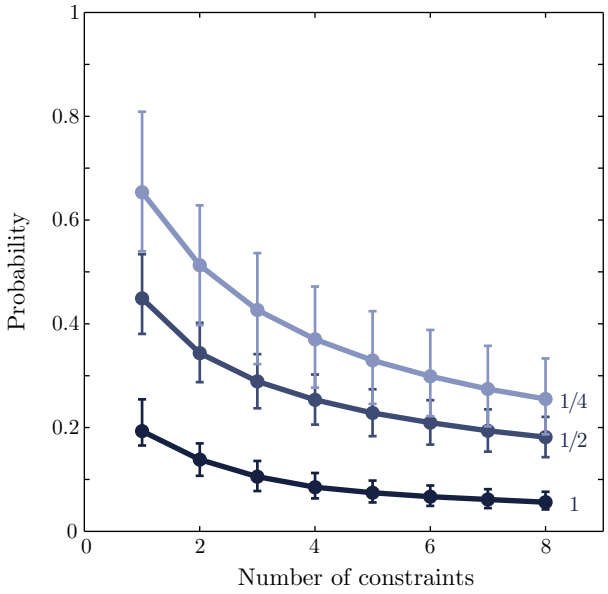


Figure 3.7: The median probability that a randomly selected light direction will satisfy between one and eight shading constraints. Each curve corresponds to constraints built from the entire illuminated portion of the occluding contour (bottom), half of the contour (middle), or a quarter of the contour (top). The error bars correspond to the 25^{th} and 75^{th} quantile. This specificity can be compared to that of shadow constraints, Figure 2.7 (left).

This probability decreased to 26% when all 8 constraints were used, Figure 3.7. Overall, the specificity is directly proportional to the number and extent of constraints. Note that this specificity can be compared to that of simulated shadow constraints, Figure 2.7 (left).

We also measured the decrease in specificity that results when the camera intrinsic parameters are unknown. Unknown camera parameters led to only a modest increase in the width of the constraints, and hence the specificity. In particular, the percentage of random light directions that satisfied constraints with unknown parameters increased by an average of only 1.8% over those with known parameters.

And lastly, recall that we exclude any concave constraints as they cannot be incorporated into the linear program. In the above simulations, an average of only 1.2% of the constraints were concave.

3.2.2 Real world

Shown in the top left panel of Figure 3.8 is an authentic photo. This scene was illuminated by a single light source located behind the camera and approximately 10 m from the scene. Shown in the top right panel are a variety of shading and shadow constraints, magnified views of which are shown beneath.

Shading constraints (black dots) are computed from segments along occluding contours (outlined in yellow). These include cloth, ceramic, and organic materials. Only the reverse constraints are shown, corresponding to the case when the light is behind the camera. These

shading constraints are physically plausible, as indicated by their feasible region (outlined in black).

Shadow constraints are specified by identifying correspondences between shadows and objects (white squares). Note that these constraints may be specified from shadow to object (labels b, c, e), or from object to shadow (labels a, d, f). As with the shading constraints, only the reverse constraints are shown. These shadow constraints are mutually physically plausible, as indicated by their feasible region (outlined in white).

Because the individual feasible regions intersect, the shading and shadow constraints are mutually consistent. This common intersection of all shading and shadow constraints correctly suggests that the lighting in this scene is consistent with a single light source.

Shown in the bottom left panel of Figure 3.8 is a composite image in which three objects (the squash and the green and orange shirts) have been replaced with versions photographed under different lighting. The same set of shading and shadows constraints were used as in the original version of this image. As shown in the lower right panel, however, the constraints are no longer satisfiable: the shadow constraints labeled a and f have a common intersection to the right, while the shading constraints labeled $8, 9, 0$ have a common intersection to the left.

For the authentic photo in Figure 3.8, a viable solution was found with a focal length $f = 60$ mm and image center $\boldsymbol{c} = (0, 0)$. For the composite photo, no combination of focal length and image center yields a viable solution.

Shown in the top left panel of Figure 3.9 is an authentic photo of a scene illuminated by the sun. Objects with distinct points were placed in the scene to allow for precise localization of the light source. Shown in the top right panel are a variety of shading and shadow constraints; only the reverse constraints are shown. The available shading constraints are physically plausible, as indicated by their feasible region outlined in black, The shadow constraints are also physically plausible as indicated by their feasible region outlined in white. Furthermore, the shading and shadow constraints are mutually consistent because the two feasible regions intersect.

Shown in the bottom left panel of Figure 3.9 is a composite image in which a pink ball has been added. The remaining cast shadows were then modified so as to be physically consistent with the inserted sphere. As seen in the lower right panel, the shading constraints are physically plausible. This is because the pink ball was photographed under lighting that was similar to the original scene. The shadow constraints are also physically plausible. Because these two feasible regions do not intersect, however, the shading and shadow constraints are seen to not be physically plausible.

Shown in top left panel of Figure 3.10 is a authentic photo that at first glance may appear unlikely[2]. Segments of approximately constant reflectance are selected along a variety of occluding contours, labeled 1-8. The outlined black region is the common intersection of

[2]Its authenticity is corroborated by additional photos on the source website.

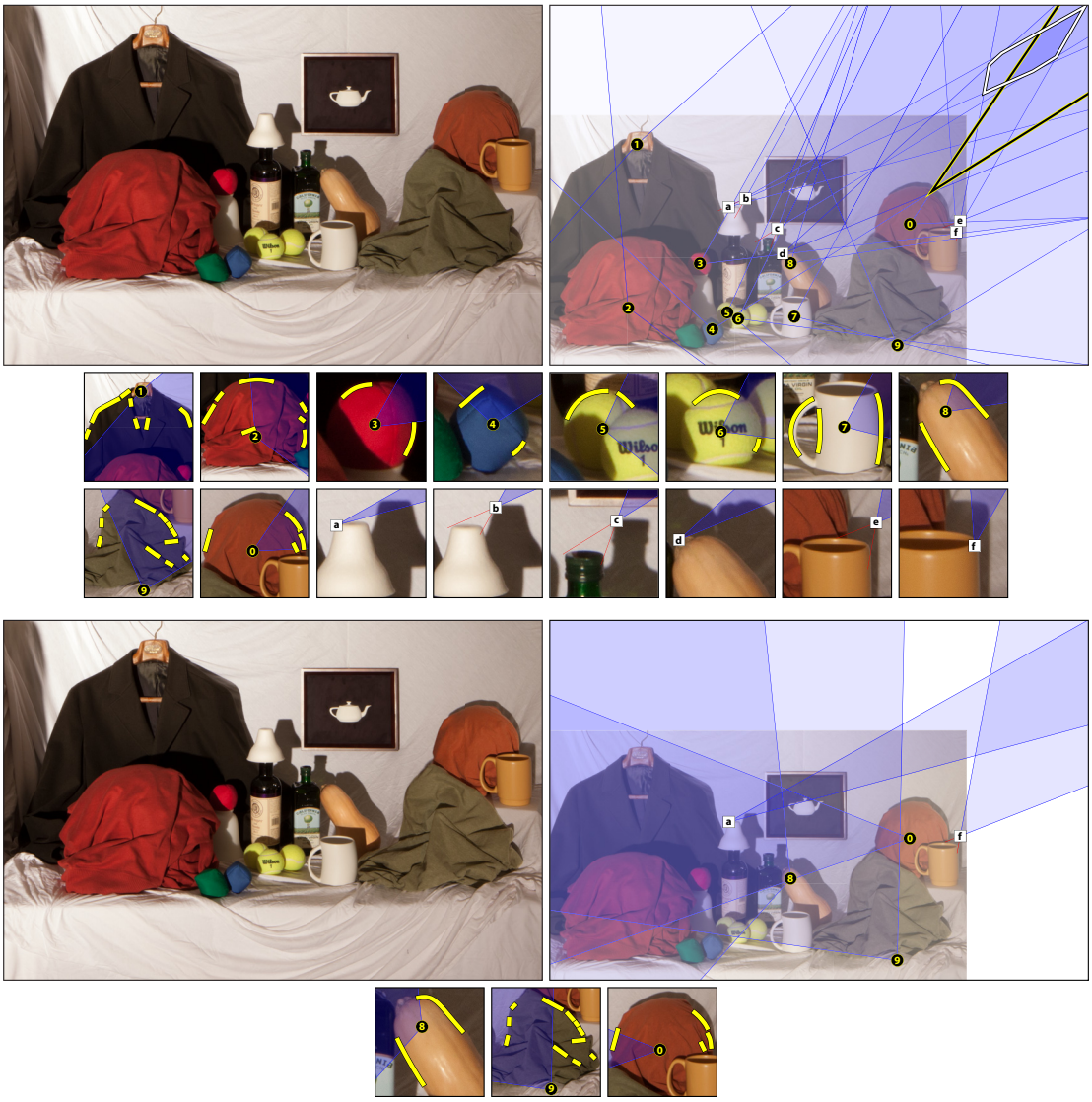


Figure 3.8: A complete shading and shadow analysis of an authentic (top) and composite (bottom) photo. Top: shading constraints, denoted by black dots, are selected from a variety of occluding contours each outlined in yellow. The outlined black region is the common intersection of these constraints, showing that they are all mutually consistent. Shadow constraints are denoted by white squares. The outlined white region is the common intersection of these constraints, showing that the shadows are mutually consistent. Because the viable shading and shadow regions intersect, the lighting in this photo is physically plausible. Bottom: in this composite photo three objects with different lighting were inserted (the squash and the green and orange shirts). The five constraints shown in the lower right are inconsistent, revealing this photo to be a composite.

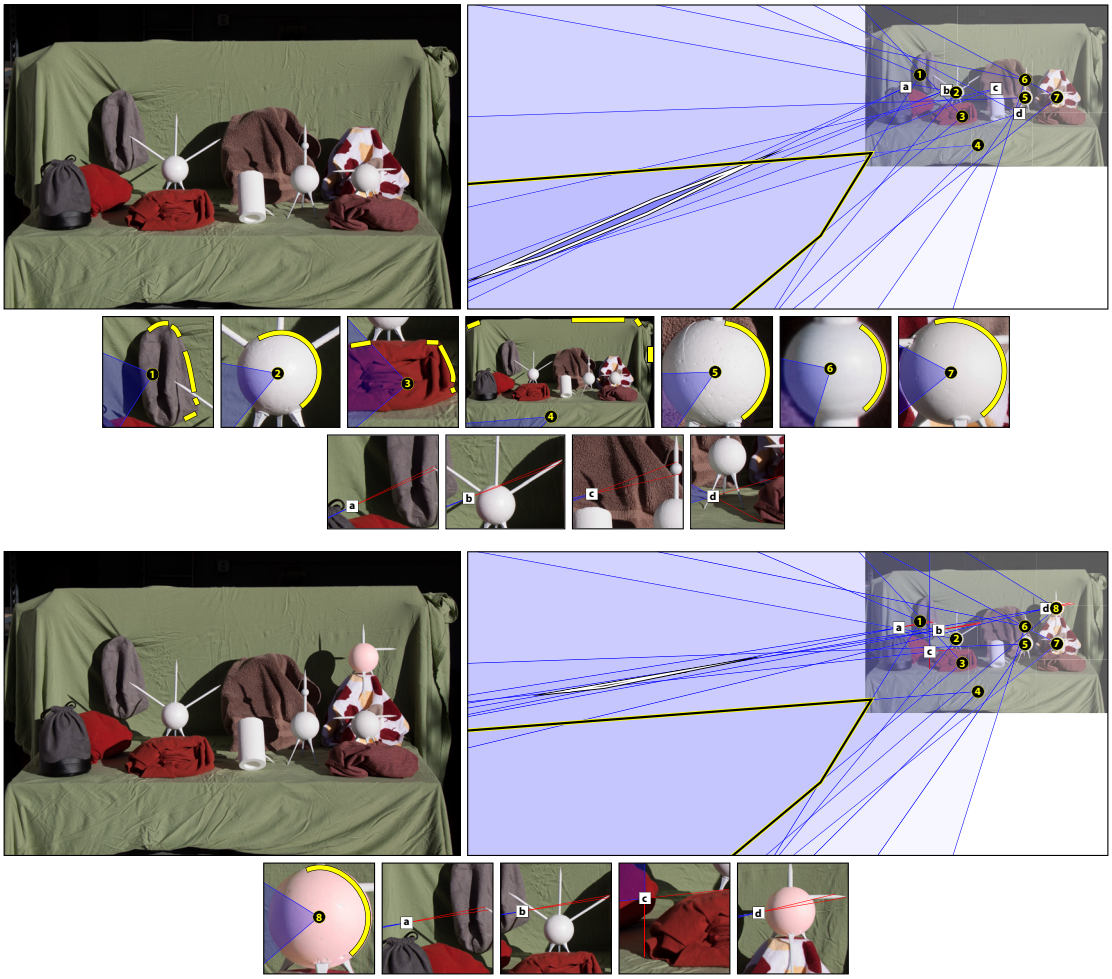


Figure 3.9: A complete shading and shadow analysis of an authentic (top) and composite (bottom) photo. Top: shading constraints, denoted by black dots, are selected from a variety of occluding contours each outlined in yellow. The outlined black region is the common intersection of these constraints, showing that they are all mutually consistent. Shadow constraints are denoted by white squares. The outlined white region is the common intersection of these constraints, showing that the shadows are mutually consistent. Because the viable shading and shadow regions intersect, the lighting in this photo is physically plausible. Bottom: the pink ball in the upper right along with its cast shadow were inserted into the image. The remaining cast shadows were then digitally altered to be mutually consistent, as seen by the outlined white region. Despite the addition of the shading constraint from the pink ball, the shading on the objects is also mutually consistent, as seen by the outlined black region. The combination of shading and shadow constraints, however, is inconsistent: the viable region of the shading does not intersect the viable region of the shadows.

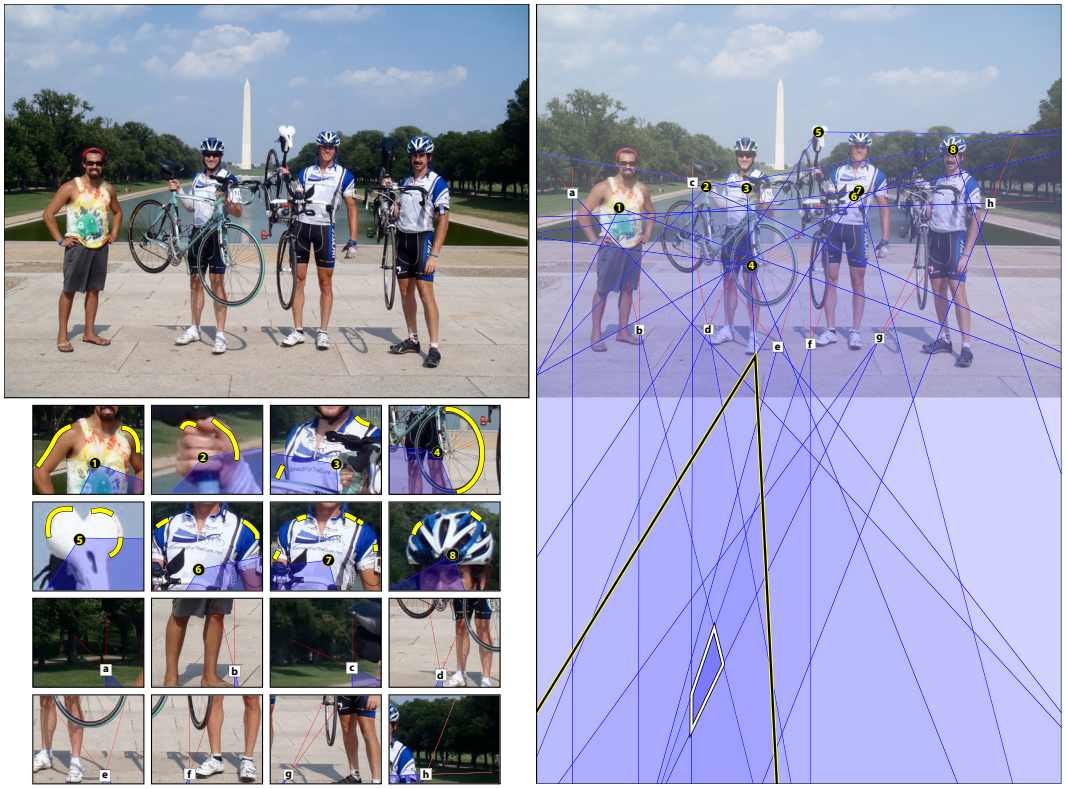


Figure 3.10: An authentic photo and the result of a shading and shadow consistency check. The region of plausible light positions that satisfy the shading constraints is outlined in black, and for the shadows, in white. These regions intersect, correctly suggesting that the shading and shadow is physically plausible.

these constraints. Shadow constraints, labeled a-g, have a common intersection outlined in white. Because the viable shading and shadow regions intersect, the lighting in this photo is physically plausible.

Shown in the top panel of Figure 3.11 is a composite photo in which the man and boy were inserted from an unrelated photo. Despite the somewhat fantastic nature of the photo, the shading and shadows appear physically plausible. Shown in the lower left are a subset of shading and shadow constraints. Shading constraints are specified along the woman's shoulders, the two hats, and along the man's hands. Cast shadow constraints are labeled a, b, d, e, and f, and attached shadows are labeled c and g. This set of shading and shadow constraints are not satisfiable. Shown in the lower right are two subsets of unsatisfiable constraints that separately entail both the man (top) and boy (bottom). There does not exist a light source in front of (shaded red) or behind the camera (shaded blue) that satisfies these constraints under any plausible camera.

Figure 3.11: A composite photo in which the man and boy were inserted. Bottom left: shading constraints are labeled 1-4, cast shadow constraints specified from shadow to object are labeled a, b, d, cast shadows specified from object to shadow are labeled e, f, and attached shadows are labeled c, g. Bottom right: two subsets of mutually inconsistent constraints revealing this image to be a composite with inconsistent lighting.

3.3 Discussion

We have developed a forensic technique for measuring the consistency, or lack thereof, of shading and cast and attached shadows. Unlike previous approaches, instead of explicitly trying to estimate the lighting in a scene, we seek to extract and combine a multitude of partial constraints on the projected location of the light source. This has the advantage that we can contend with situations in which the explicit estimation of lighting is under-constrained, a common problem in a forensic setting in which only a single image of unconstrained origin is available.

Our analysis assumes that a scene contains a single distant point light source that illuminates a Lambertian surface of constant reflectance, and is imaged under linear perspective. Although this precludes the analysis of scenes lit by multiple sources or diffuse area lights, it includes the common situation of outdoor scenes. In addition, we have shown that modest deviations from these assumptions do not severely hamper our technique. Should other violations of our assumptions (e.g., lens distortion) become an issue, they can be incorporated into our existing model of uncertainty.

We have described how the occluding contours on objects can be used to partially constrain the location of the light source because this geometric constraint is commonly available in photos. In some photos, objects may be present for which the 3-D geometry can be estimated. For example, Kee and Farid [117] describe how the known geometry of the human face can be used estimate a 3-D model, and to detect inconsistencies in 3-D lighting environments. This basic approach could be adapted to estimate the light direction, and would provide a point constraint on the location of the light source. Future work in this direction will require a model for the uncertainty in such point constraints on the location of the light.

Note that a talented forger could craft a forgery such that the shading and shadows are consistent by our measure. For example, a photo could be augmented with a synthetic object that is rendered with an approximation of the scene lighting [114]. Or an object can be re-shaded and its shadows warped to conform with the lighting. While this is not impossible it does, at a minimum, make the creation of a compelling forgery more difficult and time consuming.

Chapter 4

Applications

We have thus far developed methods to detect physically inconsistent shadows and shading in photos. In this chapter, we describe prior methods that can be generalized, by specifying partial constraints, to analyze photos in which explicit estimation of the required property is not possible. A short italicized abstract is provided with each method for readers who wish to quickly survey these applications.

4.1 Center of projection

The method of O'Brien and Farid [186] detects inconsistent centers of projection among objects in a photo. To estimate the center of projection from an object, three mutually orthogonal vanishing points are required, and O'Brien and Farid find them in reflecting surfaces. If the reflecting surface is rectangular, its sides specify two orthogonal vanishing points. A third orthogonal vanishing point can be found when a reflection appears: lines connecting points on an object to its reflection converge to a vanishing point. The center of projection can be estimated from these three orthogonal vanishing points. When two or more reflecting surfaces appear, differences in the centers of projection evidence forgery. Partial constraints can be used to relax the requirement that three mutually orthogonal vanishing points are present, and thus a reflection need not appear. A pair of mutually orthogonal vanishing points partially constrain the center of projection (detailed below), and any rectangle in an image therefore provides a constraint. Furthermore, O'Brien and Farid manually estimate corresponding points between an object and its reflection. Such correspondences can be difficult to specify in practice when reflections are faint or ambiguous, but can be partially constrained as well.

4.1.1 Estimation

The center of projection of a camera can be estimated by identifying vanishing points in an image, Figure 4.1 (left). Parallel rays in the world, such as those identified by the edges of the cube, converge in the image (black lines) to identify vanishing points (yellow dots).

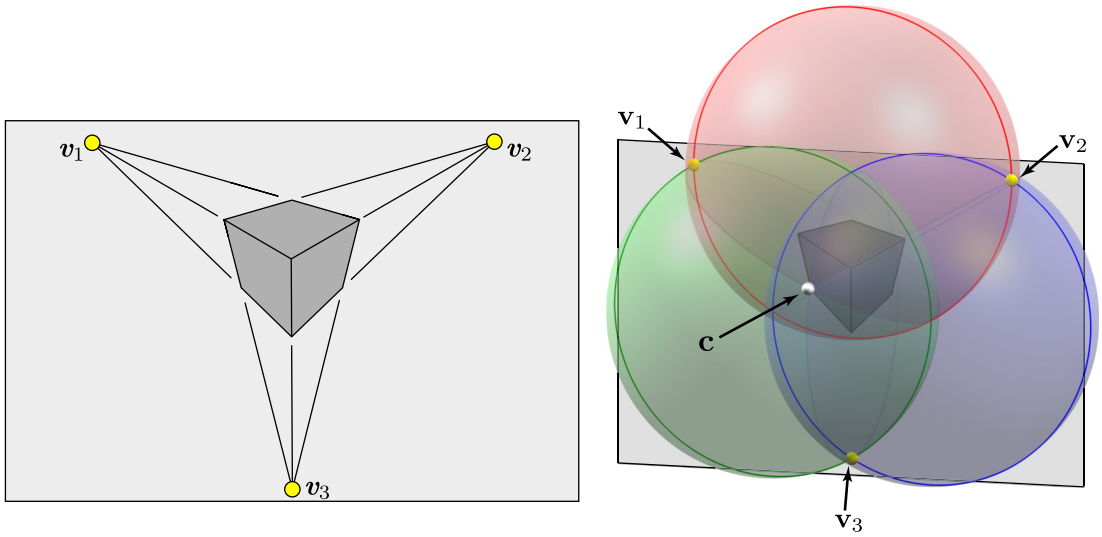


Figure 4.1: Estimating the center of projection. Parallel lines in the world intersect under perspective projection (left). Shown are three sets of parallel lines that converge to three vanishing points (yellow dots), and identify three mutually orthogonal directions in the world. The 3-D center of projection can be estimated from such 2-D vanishing points (right). Each vanishing point is placed on a plane in 3-D (yellow dots, $\mathbf{v}$). Rays through each vanishing point and the center of projection (white dot, $\mathbf{c}$) must be parallel to the edges of the cube, and are therefore mutually orthogonal. Each pair of rays specifies a quadratic constraint on the center of projection (e.g., Equation 4.1), which can be visualized as a sphere (red, blue, green). The center of projection lies at the intersection of these spheres. Figure due to O'Brien and Farid [186].

Furthermore, all rays in the world that are parallel to these edges of the cube converge to the same vanishing points: $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, and $\boldsymbol{v}_3$. The center of projection can be estimated if these vanishing points correspond to mutually orthogonal lines in the world, as is the case for those identified on the cube.

To estimate the center of projection, the 2-D vanishing points, $\boldsymbol{v}$, are placed at 3-D points $\mathbf{v}$ on an arbitrary plane in the world, Figure 4.1 (right). Rays through points $\mathbf{v}$ and the center of projection, $\mathbf{c}$ (white dot), must be parallel to the edges of the cube, and are therefore mutually orthogonal. The following constraints can therefore be placed on $\mathbf{c}$:

$$(\mathbf{c} - \mathbf{v}_1) \cdot (\mathbf{c} - \mathbf{v}_2) = 0 \tag{4.1}$$

$$(\mathbf{c} - \mathbf{v}_1) \cdot (\mathbf{c} - \mathbf{v}_3) = 0 \tag{4.2}$$

$$(\mathbf{c} - \mathbf{v}_2) \cdot (\mathbf{c} - \mathbf{v}_3) = 0 \ . \tag{4.3}$$

This quadratic system of equations can be solved to estimate the center of projection [186]. Intuitively, each quadratic constraint specifies that the center of projection must lie on a sphere in the world, Figure 4.1 (red, green, and blue spheres). The intersection of these spheres specifies the center of projection.

O'Brien and Farid [186] suggest a clever way to identify three vanishing points that can

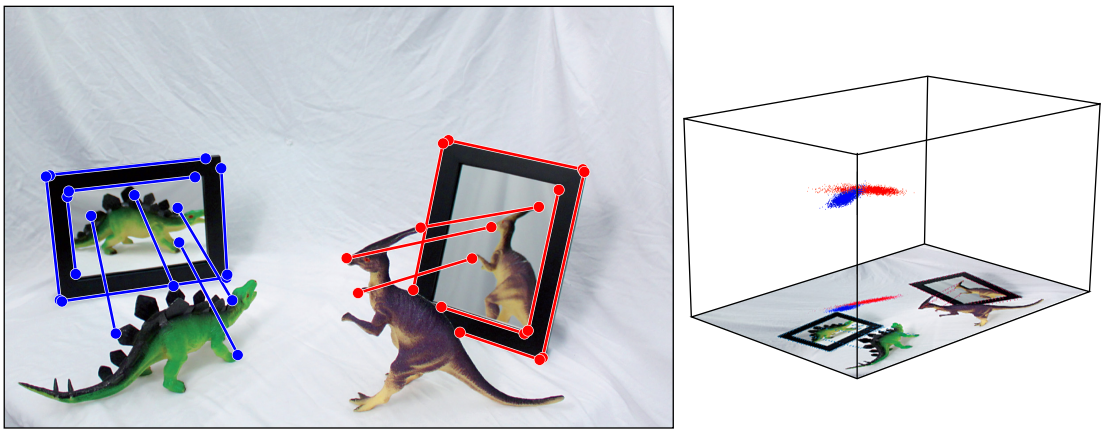


Figure 4.2: Determining the consistency of the center of projection using the estimation approach. Lines are specified along the edges of reflecting surfaces (left), and between points on an object and its reflection. Such lines correspond to three mutually orthogonal directions in the world [186]. The center of projection can therefore be independently estimated for two reflecting surfaces (red, blue) by solving Equations (4.1) - (4.3). This solution is sensitive to small changes in the lines, and a perturbation analysis is performed (right). Specifically, the endpoints of each line are repeatedly perturbed and the center of projection is estimated. Shown are two clouds of possible centers of projection. These clouds overlap, thus correctly determining that the photo is physically consistent under a single center of projection. Figure due to O'Brien and Farid [186].

be used to estimate the center of projection. Specifically, if a reflecting surface such as a window or mirror is present in a scene, lines that connect corresponding points on the object and its reflection identify parallel rays in the world, Figure 4.2 (top left). Furthermore, many reflecting surfaces like windows have rectangular borders around them. By combining these rectangular borders with lines between an object and its reflection, three mutually orthogonal vanishing points can be identified, and the center of projection can be estimated.

The authenticity of a photo can therefore be tested using an estimation approach, provided that reflecting surfaces are present in a photo [186]. Shown in Figure 4.2 is an authentic photo that contains two objects reflected in two surfaces (left). The center of projection can be estimated for each of these reflected objects by selecting the borders of the frame and points in the reflection (blue and red; lines and dots). In practice this computation can be sensitive to the particular pixel that was selected, and a perturbation analysis is performed (right). Shown are two clouds of possible locations (blue and red) for the center of projection as estimated from the two mirrors. These clouds overlap, correctly suggesting that the photo is authentic. Forgery is detected if the distance between the nearest pair of red and blue points exceeds a threshold [186].

4.1.2 Partial constraints

We now describe how inconsistent centers of projection might be detected using partial constraints. In practice rectangular surfaces are common, but do not contain reflections

or other cues that specify a third, orthogonal, vanishing point that is needed to estimate the center of projection. Rectangular surfaces do however specify two orthogonal vanishing points, which partially constrain the center of projection to a sphere, Figure 4.1 (right). If a purportedly authentic photo contains two or more rectangular surfaces, such as the faces of the cube, the corresponding spheres must intersect (e.g., red and blue).

A simple way to determine if two such spherical constraints intersect is to consider their circular intersection with the image plane, Figure 4.1 (red and blue circles). These spheres are bisected by the image plane. To see this note that a pair of vanishing points uniquely determine each sphere, as in Equation (4.1), and lie on its surface. Furthermore, these vanishing points must lie on a line that specifies the diameter of the sphere; if not, the sphere would not be uniquely determined. This diameter line, which connects the vanishing points and lies in the image plane, must pass through the center of the sphere. The sphere must therefore be bisected by the image plane, and its intersection forms a great circle.

Because each spherical constraint is bisected by the image plane, pairs of spherical constraints can only intersect if their corresponding great circles do. This provides a simple test for the physical plausibility of two pairs of vanishing points. Shown in Figure 4.3 is an example of this partial constraint. An authentic photo (top left) that does not contain an object and its reflection was emulated by removing the dinosaurs from the authentic image in Figure 4.2 (left). Although three orthogonal vanishing points cannot be identified, each frame specifies a pair of orthogonal vanishing points that constrain the center of projection to a sphere. Shown on the right are the associated great circles (red and blue circles), which can be drawn through the vanishing points without explicitly computing the intersection of the sphere and the image plane. These great circles are very large, and the photo area is visible within a small black rectangle. Because the circles overlap, the photo is correctly determined to be authentic.

A falsified photo is also shown in which the rectangular frames have been warped, Figure 4.3 (bottom left). Although the modification is visually subtle, the circular constraints do not overlap (bottom right, red and blue). Furthermore, a perturbation analysis similar to that performed by O'Brien and Farid does not produce any pair of circles that intersect. We therefore correctly determine that the objects do not specify a physically plausible center of projection.

A partial constraint method might therefore be used to determine if multiple rectangular regions are physically plausible. Circles can be constructed for each rectangular region, and a photo is implausible if any pair does not intersect. The perturbation analysis proposed by O'Brien and Farid might furthermore be simplified by allowing the analyst to partially constrain the orientation of the lines. Specifically, the analyst would select a first endpoint and a range of visually plausible second endpoints. This would specify a wedge-shaped region of plausible lines along an edge of the rectangular surface. The intersection of multiple wedges would specify a region of plausible vanishing points that model the uncertainty in

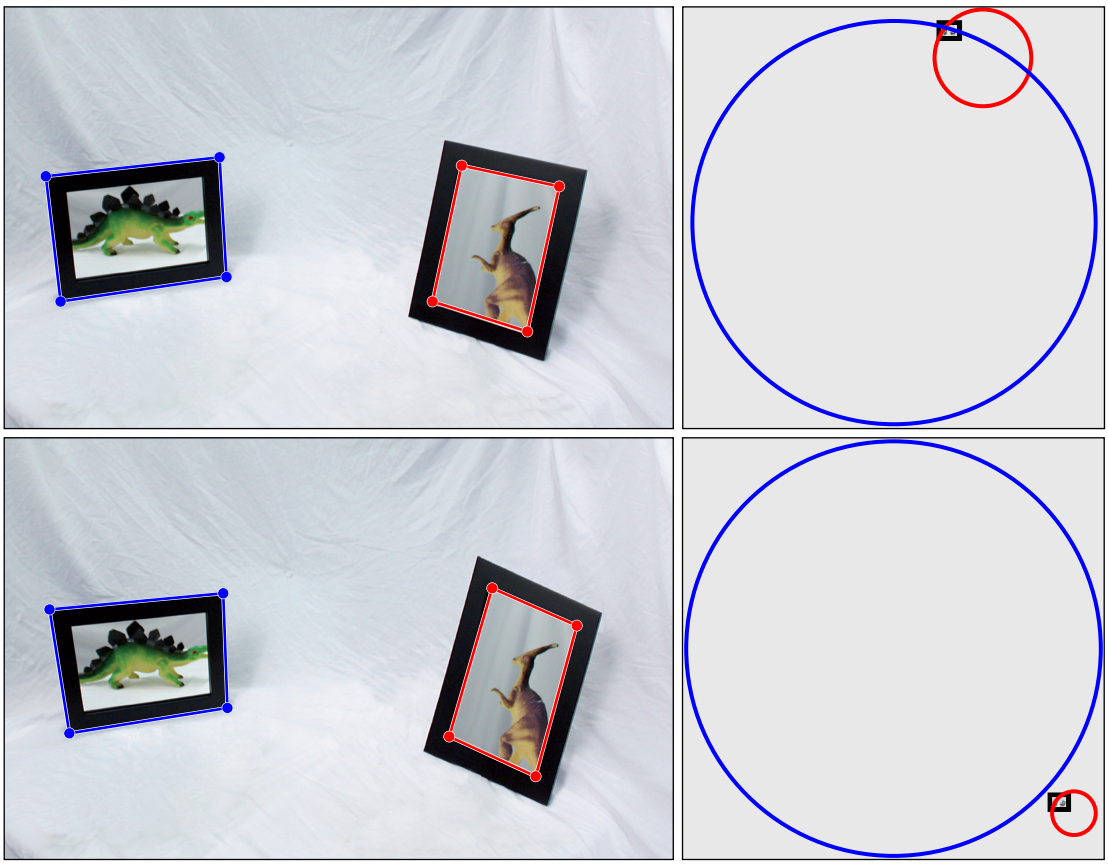


Figure 4.3: Determining the consistency of the center of projection using partial constraints. In practice, reflections are often not present in photos (top left), but are required by O'Brien and Farid [186] to estimate the center of projection. Partial constraints on the center of projection can however be specified using the pair of vanishing points that is provided by the edges of the frames (blue, red). Specifically, the center of projection can be constrained to a sphere for each frame. Each sphere is bisected by the image plane, and specifies a circle (top right). (See also Figure 4.1.) A plausible center of projection exists if the spheres, and thus the circles, intersect. Shown in the bottom row is a falsified photo in which the frames have been warped. Although this change is visually subtle, the circular constraints specified by the frames (red, blue) do not intersect (right). Furthermore, a perturbation analysis does not yield overlapping circles. We therefore correctly determine that the photo has been falsified. Photo content is due to O'Brien and Farid [186], and is modified here for illustration.

the image content.

Although the circles provide a partial constraint, note that this simple approach requires binary comparisons to detect forgery, and will therefore overlook inconsistencies that entail three or more surfaces. A fully general approach will specify a full spherical constraint on 3-D center of projection from each rectangular surface.

4.1.3 Comparison

The approach of O'Brien and Farid [186] faces challenges that result from the estimation approach. Foremost is the somewhat restrictive assumption that a rectangular surface contain a reflection. This assumption is necessary to estimate the center of projection, as the reflection specifies the necessary third vanishing point, but it restricts the photos to which the method can be applied. In contrast, the suggested partial constraint method could be applied to any region of a photo that specifies two orthogonal vanishing points (e.g., a rectangular surface). Because these surfaces are more common, multiple partial constraints may be available in an image, and may be mutually inconsistent. The combined consistency of multiple surfaces can be determined by specifying a spherical constraint on the 3-D center of projection.

The requirement that corresponding points between an object and its reflection be explicitly estimated is also somewhat restrictive. O'Brien and Farid use these correspondences to estimate and detect inconsistent centers of projection. They also use these correspondences to detect forgery when multiple objects are reflected in the same surface (not described above). All objects that are reflected in the same surface must identify lines which converge to the same vanishing point [186]. O'Brien and Farid assume that these lines can be estimated from corresponding points, but these points can be difficult to specify if a reflection is faint or the object does not have distinct features.

Partial constraints can be used to relax the requirement that corresponding points be explicitly estimated. Specifically, we describe in Chapter 2 how wedge-shaped constraints can be used partially constrain manually-specified shadow-to-object correspondences, and to constrain the location of a point light source. The implementation of the method in Chapter 2 can therefore be used, without modification, to analyze multiple reflections that appear in the same surface when correspondences cannot be estimated.

4.2 Chromatic aberration

The method of Johnson and Farid [106] detects inconsistent patterns of chromatic aberration in photos. Specifically, inconsistent lateral chromatic aberrations are detected, which manifest in photos as a misalignment between the color channels, Figure 4.4. To detect inconsistent aberrations, a global alignment is first computed. This global vector field is then compared to local, block-wise, estimates. Forgery can be detected if the angle between the vectors in the global and local estimates is large. Partial constraints can be used to relax the requirement that a block contain sufficient image content to estimate the chromatic aberration pattern. Each vector constrains the center of the aberration, and its direction can be partially constrained from a single edge that exhibits chromatic aberration (detailed below). Such constraints can be manually or automatically specified at a sparse set of edges, thereby enabling the analysis of regions in which image content is insufficient for estimation.

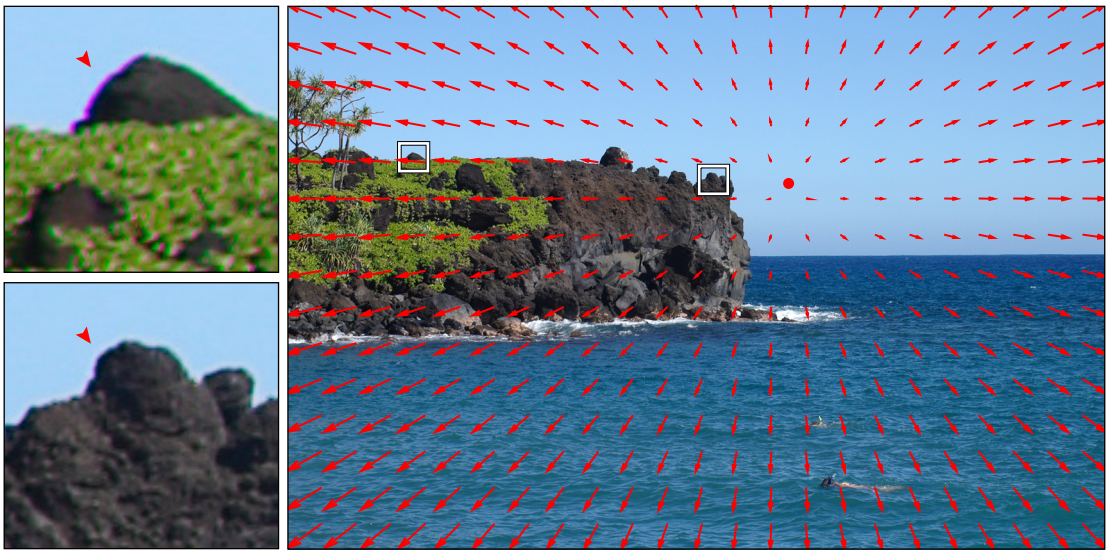


Figure 4.4: Lateral chromatic aberration can be described as a misalignment of the color channels with respect to one another. This distortion creates unnatural colors that are most salient along the edges of objects (left, top). The visual impact of the distortion varies across the image (left, bottom), and forms a pattern across the image (right). Shown is a vector field that aligns the blue and green channels of the photo; vectors are scaled by a factor of 150 for display. This field can be described as a radial expansion by a scalar factor α about a central point $\boldsymbol{c}$ (red dot). Figure due to Johnson and Farid [106].

4.2.1 Estimation

Lateral chromatic aberration creates unnatural colors along edges in photos, Figure 4.4 (left), and varies within a photo (top left, bottom left). These aberrations result from a global misalignment between color channels that can be modeled as an expansion or contraction about a central point (right). Shown is a vector field which visualizes an expansion of the green channel of the photo, and is described by its center $\boldsymbol{c}$ and expansion factor α:

$$\boldsymbol{p} \;=\; \alpha(\boldsymbol{q}-\boldsymbol{c})+\boldsymbol{c} \;\;, \tag{4.4}$$

where vector $\boldsymbol{q}$ is a point in one color channel, and $\boldsymbol{p}$ its corresponding location in another channel. The parameters of this vector field, $\boldsymbol{c}$ and α, can be estimated by aligning the requisite pair of color channels (in this case the green and red channels). Johnson and Farid perform this alignment by optimizing mutual information [106].

Forgery can be detected by following an estimation approach [106]. First a global aberration field is estimated for the photo. (The authors assume that falsified regions are small and will not significantly affect this global estimate.) Next, the photo is divided into overlapping blocks and a local distortion field is estimated for each block. Forgery is detected by computing differences between the block and global distortion fields. Specifically, the angle between each vector in the block and global distortion field is computed. If the average

angle is above 60°, there is a 97% chance that the region has been altered [106].

4.2.2 Partial constraints

We now describe how inconsistent chromatic aberrations might be detected using a partial constraint approach. To begin, suppose that the vector field that aligns a pair of color channels in a potentially-falsified photo is known, Figure 4.4. Each vector constrains the chromatic aberration parameters. By visual inspection, it can be seen that the position and direction of each vector constrains the center of the aberration, $\boldsymbol{c}$ (dot), to lie on a line, Equation 4.4. The length of each vector constrains α. Because Johnson and Farid consider only the vector direction, we will however omit discussion of α for simplicity.

In practice it can be difficult to estimate the alignment vectors that describe chromatic aberration. For example, Johnson and Farid note that estimation is poorly constrained when image texture is limited [106]. Furthermore, chromatic aberration is sensitive to resizing and compression. A sparse set of edges that exhibit chromatic aberration might nonetheless be present in a photo.

Although a single edge is not sufficient to estimate its alignment vector, note that the chromatic aberration along the edge can be removed by simply translating the color channels in the local region. This translation partially constrains the alignment vector to lie on one side of the edge. (To see this, note that the alignment vector is constrained only in the direction that is orthogonal to the edge, due to the aperture problem.) Furthermore, this partial constraint can be constructed without explicitly estimating the translation: we need only determine which of two directions, orthogonal to the edge, will reduce the aberration. This relaxed problem may be more computationally straightforward than explicit estimation, and it can also be performed manually by an analyst.

Given a partial constraint on the alignment vector at an edge, we would like to constrain the center of the aberration, $\boldsymbol{c}$. If $\boldsymbol{c}$ is known to lie behind the alignment vector, as illustrated in Figure 4.4, then it must lie within a half-plane that is on the opposite side of the edge. In practice, however, the aberration center may lie in front of, or behind, an alignment vector because aberrations may entail a contraction or expansion of a color channel. A single edge cannot therefore constrain the aberration center. If multiple edges are present, however, the center must lie in front of, or behind, all alignment vectors and thus all half-plane constraints. This provides a simple forensic test for the physical plausibility of chromatic aberrations. The consistency of multiple half-plane constraints can be determined by constructing a linear program like that used for shadows, Section 2.1.3. Like shadow constraints, two linear programs are constructed, one with all half-plane constraints inverted, and the chromatic aberration is physically plausible if either is satisfiable.

Shown in Figure 4.5 is an example of an analysis of chromatic aberration in which partial constraints were specified manually. A falsified photo was created by adding a shark fin (lower left) to an original photo provided by Johnson and Farid [106]. Differing lateral

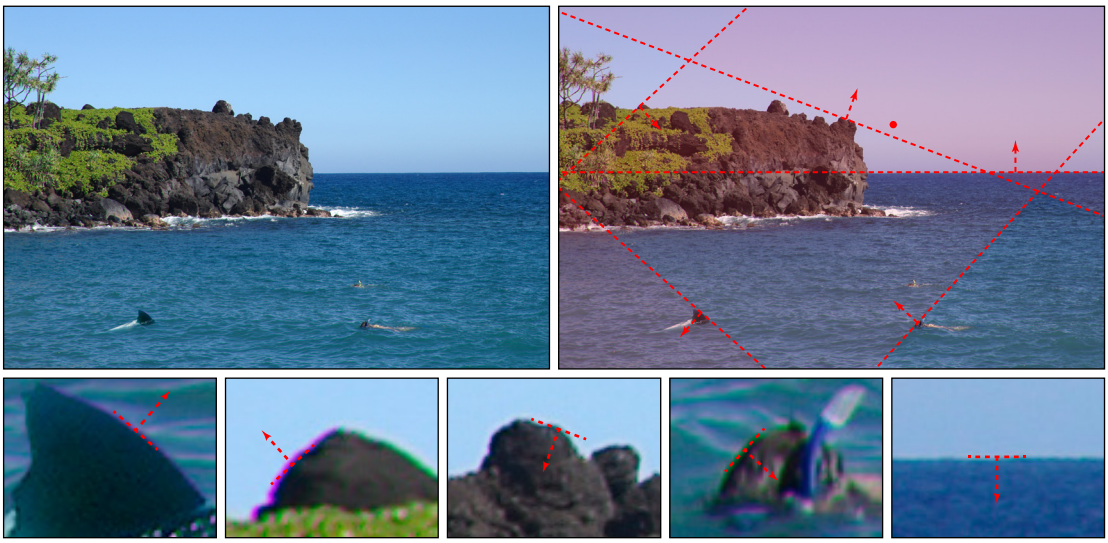


Figure 4.5: Detecting inconsistent chromatic aberration with partial constraints. An original photo was falsified by adding a shark fin (top left), for which a single edge (bottom left detail) has chromatic aberration that is inconsistent with the remainder of the photo. A single edge is insufficient to estimate chromatic aberration, but partial constraints can be used to detect the forgery. An observer manually translated the green channel of the photo to locally reduce the severity of the aberration on five edges (bottom panels); shown are the translation directions (dashed vectors). These translations specify half-plane constraints, on the center of the aberration, that must be satisfied as input, or in their inverse directions (top right). The constraints are not satisfiable, thus correctly suggesting that the aberration on the shark fin is not physically plausible. Photo due to Johnson and Farid [106].

chromatic aberration was introduced on the front edge of the fin, and in the background (emphasized for display). Because a single edge is insufficient to estimate chromatic aberration, it would be difficult to detect this forgery using the method of Johnson and Farid.

An observer manually analyzed the chromatic aberration by adjusting each of five edges (bottom insets) to locally reduce the severity of the chromatic aberration. The observer visually determined that the green channel needed to be translated in the illustrated directions (dashed vectors), orthogonal to the edge (dashed segments). These translations specified five half-plane constraints from which two linear programs were constructed. In the first linear program, half-planes were oriented as input by the analyst; in the second, all half-planes were inverted. Both linear programs are unsatisfiable. For example, when all constraints are inverted (top right), the half-plane on the shark fin does not intersect the feasible region that would be defined if the remaining for constraints were considered alone. Furthermore, the feasible region specified by the authentic edges correctly contains the center of the aberration (dot). We therefore correctly determine that the chromatic aberration on the edge of the shark fin is not physically plausible with the other elements in the image.

4.2.3 Comparison

The approach of Johnson and Farid [106] faces challenges that are characteristic of an estimation approach. Blocks must be sufficiently large to estimate the aberration (300×300 blocks are used), which restricts the minimum size of the forgery that can be detected. The authors also note that estimation of the aberration parameters can be difficult if a block contains little or no spatial frequency content (e.g., a region such as the sky), and blocks are therefore used only if their average gradient magnitude is high. Lastly, chromatic aberration is sensitive to resizing and compression, which may leave behind few edges that exhibit such distortion. In contrast, partial constraints can be used to detect inconsistent chromatic aberrations from isolated edges, which cannot be used to estimate the aberration vector. A partial constraint can be specified by determining which of two translation directions, orthogonal to the edge, will reduce the aberration. This relaxed problem can be solved manually if necessary.

4.3 Reflection

The method of Johnson and Farid [108] detects specular reflections that are inconsistent with a point light source. This detection requires that the 3-D surface normal at the specular point be estimated. Although the 3-D normal is typically unknown, Johnson and Farid note that specular reflections often appear in peoples' eyes. The known geometry of the human eye is therefore used to estimate the 3-D surface normal. Once the surface normal is known, it is possible to estimate the 3-D direction to the point light source. This process is repeated for each person in a photo, and inconsistent light directions between subjects evidence forgery. Partial constraints can be used to relax the assumption that specular reflections appear in peoples' eyes. The 3-D surface normal can be partially constrained by an analyst (detailed below), and specular reflections can be analyzed without explicit knowledge of the geometry of the reflecting object.

4.3.1 Estimation

Specular reflections can be used to estimate the light direction by a simple geometric construction, Figure 4.6 (left). Shown is an image plane containing an image of a sphere; the camera center of projection is denoted by ◉. A specular reflection appears in the image at a point (center of white spot) where the surface normal (solid vector) is midway between the camera ray and the light direction (dashed vector). More precisely, the angles between the surface normal and both the camera ray and light direction are equivalent (black hatched arcs). It is therefore possible to estimate the light direction if both the surface normal and the camera ray are known.

Johnson and Farid [108] developed a method by which the surface normal and camera ray can be estimated at a specular reflection. Specifically, if a photo contains people, specular

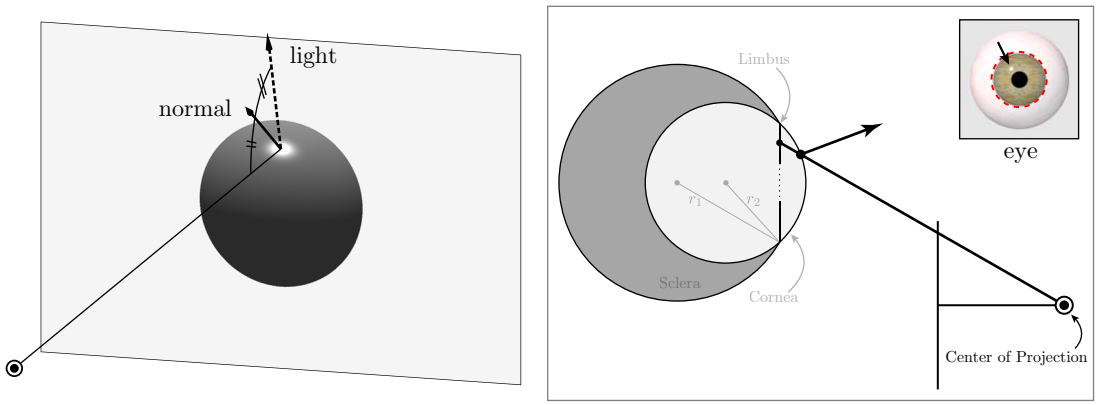


Figure 4.6: Determining the consistency of specular reflections by estimation [108]. **Left:** shown is an image plane containing an image of a sphere; the camera center of projection is denoted ◉. A specular reflection appears where the sphere's surface normal (solid vector) is halfway between the light direction (dashed vector) and the camera ray (solid line). The light direction can be estimated if the surface normal and center of projection are known. **Right:** the surface normal (solid vector) and center of projection can be estimated from specular reflections in people's eyes (inset, arrow) [108]. If multiple people appear in a photo, and have specular reflections in their eyes, light directions can be estimated for each subject and compared for consistency.

reflections often appear in their eyes, Figure 4.6 (right). Inset is a synthetically rendered eye that contains such a reflection (black arrow). The surface normal at this specularity can be estimated from the known geometry of the eye, which comprises two spheres of known radii (schematic, dark and light gray circles). The surface normal is found by drawing a ray from the center of projection through the image of the specularity, and to the smaller sphere. This construction requires that the camera's center of projection and the 3-D orientation of the eye are known. Johnson and Farid estimate this information from the circular boundary of the eye (inset, dashed red). This circular boundary typically images as an ellipse under perspective projection. A homography can be estimated which maps the ellipse back to a circle; the camera parameters and orientation of the eye can be estimated by factoring this homography [108].

Forgery can therefore be detected by estimating the light direction from specular reflections in peoples' eyes. If a photo contains multiple people, a light direction can be estimated for each person. If the light direction is purportedly the same among the subjects, angular differences between the estimated light directions that exceed 10° to 15° evidence forgery [108].

4.3.2 Partial constraints

A partial constraint approach can be used to relax the requirement that specular points appear in subjects' eyes. To begin, recall the geometry of specular reflection, Figure 4.7 (top left). Shown is an image plane that contains an image of a sphere. If the camera ray (solid line) and surface normal are known at a specular point, the lighting vector can

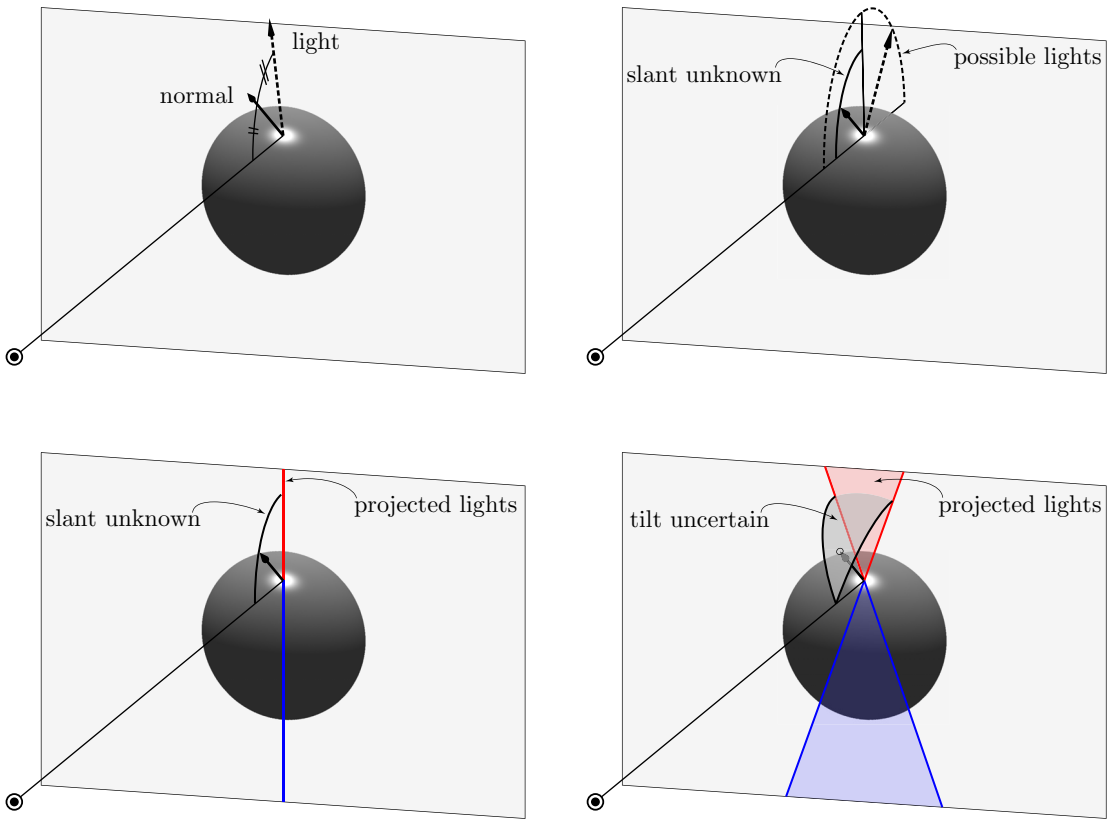


Figure 4.7: Partial constraints from specular points. **Top left:** shown is an image plane containing an image of a sphere; the camera center of projection is denoted ◉. A specular reflection appears where the sphere's surface normal (solid vector) is halfway between the light direction (dashed vector) and the camera ray (solid line). The surface normal can be simply rotated away from the camera ray to construct the light vector; such rotations change the slant of the vector. **Top right:** the slant of the surface normal is typically unknown, but lies between zero and 90 degrees. The slant of the light vector, which is double that of the normal, must therefore lie between zero and 180 degrees. **Bottom left:** the projected location of distant light sources can be partially constrained to a line in the image, which identifies lights that are in front of the camera (red segment), and behind the camera (blue segment). This constraint assumes that the direction of the slant, or tilt, of the surface normal is known. **Bottom right:** when the tilt of the surface normal is uncertain, multiple line constraints may be plausible. Such lines form a wedge-shaped region in the image that partially constrains the location of the light (red and blue shading).

be constructed by simply rotating the surface normal away from the camera ray. Such rotations change the so-called slant of the surface normal. The slant angle is typically unknown, Figure 4.7 (top right). Plausible slants (solid arc) lie between zero degrees (by definition) and 90 degrees, because the surface cannot face away from the camera. The slant of the lighting vector therefore ranges between zero and 180 degrees (dashed arc).

A partial constraint can be defined on the projected location of the light in the image plane. Assume for a moment that the direction in which the surface normal is slanted is

known. In this case, the plausible light directions span a plane in 3-D, and distant lights will project onto a line in the image. This line comprises two segments, which correspond to light directions in front of the camera (red), and directions behind the camera that are inverted by projection (blue).[1]

In practice, the direction in which the surface normal is slanted is also typically unknown, and a simple line constraint cannot be specified on the projected location of the light. When this so-called tilt direction is uncertain, multiple line constraints may be plausible (bottom right). All such lines pass through the specular point, and form a wedge-shaped region in the image plane that partially constrains the location of the light (red and blue shading).

Although the slant of the surface normal is typically unknown, the tilt might be partially constrained by a forensic analyst. Intuitively, the tilt specifies a direction in which the surface is most steeply slanted away from the camera. Observers have demonstrated the ability to specify the tilt with precision well below $\pm 90°$, and in some cases as low as $\pm 10°$ [73, 173, 132]. In a forensic setting, a conservative approach can be taken in which the analyst specifies a range of possible tilt directions that can be logically justified or debated, much like the analysis of cast shadows in Chapter 2. This provides a partial constraint on the tilt, and therefore a wedge-shaped partial constraint on the projected location of the light source that creates a specular reflection.

Shown in Figure 4.8 is an example of how a forensic analysis of specularities might be conducted. A scene (top left) was populated with specular copper blobs of random shape, and was rendered under a distant point light source. Specular points are visible on the surface of each blob (white dots), and provide partial constraints on the projected location of the point light (top right). Each constraint is notated with a red wedge-shaped region that specifies a range of tilt directions in which the surface normal may be slanted. The inverted constraints, which are notated by blue in Figure 4.7, are hidden for clarity. The intersection (red) of the five constraints correctly contains the projected location of the point light source (yellow dot).

A falsified photo was also constructed by modifying the shading and specularity on the two left blobs, Figure 4.8 (bottom left). The specularity on the top left blob is consistent with a light source from above, and the leftmost blob was rotated. Despite these modifications, it is somewhat difficult to visually detect the inconsistency. Five constraints were specified (bottom right), and are inconsistent. The constraint from the top left blob does not intersect the other constraints, as the specularity appears at a point where the surface is sloped up and to the left, rather than toward the right. The inverted constraints are hidden for clarity, and do not intersect. The specular points in the photo are therefore correctly determined to be physically implausible under the assumption of a single distant light source. Each constraint can be debated to logically verify or challenge this analysis

[1] Johnson and Farid consider 3-D light directions, which are equivalent to constraints in the image that have forward-backward polarity (if the center of projection is known), Chapter 3.

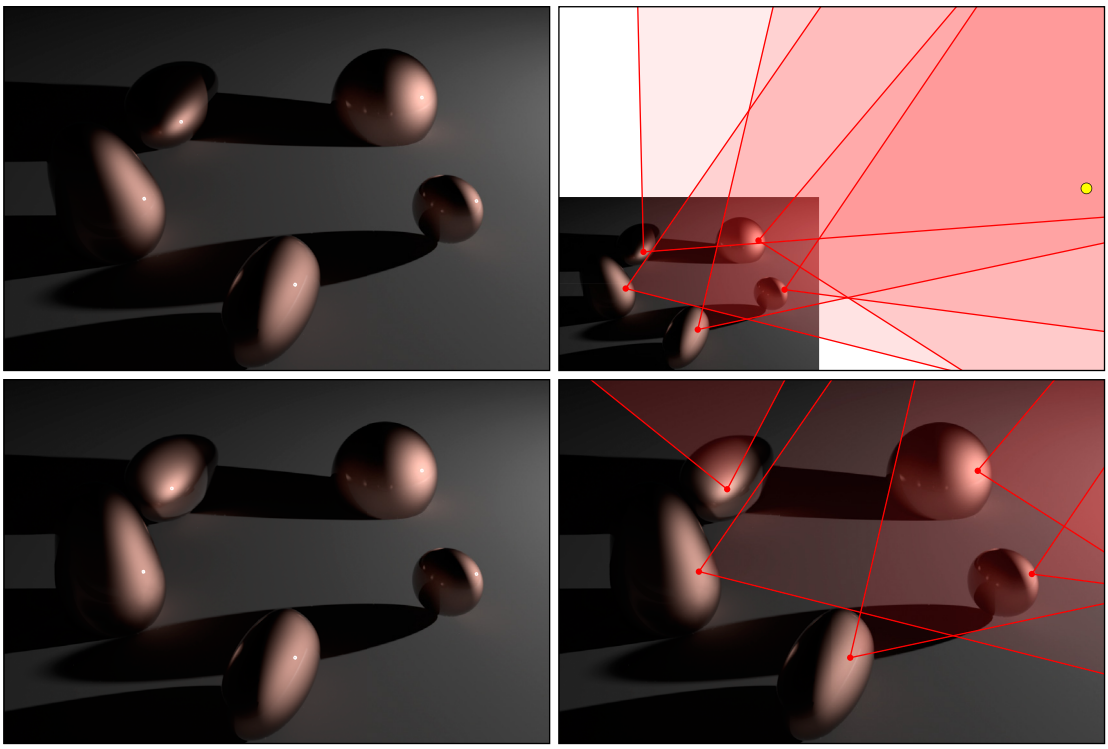


Figure 4.8: Determining the consistency of specular reflections using partial constraints. **Top left:** shown is a scene populated by specular copper blobs, and rendered under a point light source. Specular points are visible on each blob (white dots), and constrain the projected location of the point light source. **Top right:** partial constraints are manually identified by an analyst (red shading). A conservative approach is taken in which an analyst specifies a range of possible tilts (red lines) that she believes can be objectively defended. These five partial constraints correctly encompass the projected location of the light source (yellow dot). **Bottom left:** a falsified photo was created in which the left two blobs are modified. **Bottom right:** partial constraints are specified at specular points in the falsified photo. These constraints do not intersect, thus correctly evidencing forgery (the inverted blue constraints are not shown for clarity). Each constraint is debated to logically verify or challenge the analysis.

— a far more objective task than a casual visual assessment of the consistency of specular reflections (bottom left).

4.3.3 Comparison

The approach of Johnson and Farid [108] poses challenges that are characteristic of an estimation approach. Foremost is the restriction that specular points must appear in peoples' eyes, which is required to estimate the surface normal and the camera's center of projection. In contrast, a partial constraint method can be used to detect inconsistent specular reflections when the surface geometry is only partially known (i.e., by visual analysis), and when the center of projection of the camera is unknown.

Chapter 5

Discussion

The field of photo forensics has matured significantly, as methods have been developed to detect a variety of digital and physical properties that are perturbed by forgery. There exist, however, many photos in which the required property cannot be estimated. We therefore suggested an alternative approach in which properties are partially constrained, and inconsistent constraints are used as evidence. By relaxing the estimation requirement, forensic methods can be applied to a broader range of photos.

In this dissertation, we developed complimentary physically-based methods to analyze shadows and shading in photos for which it is not possible to explicitly estimate the projected location of the light source. We then described a variety of prior methods that can be generalized to detect forgery in photos for which it is not possible to explicitly estimate the required properties.

In Chapter 2, we developed methods to determine the physical consistency of cast and attached shadows that form under both point and area lights. To do so, we exploited the well known constraint that a light source must lie along a ray that connects a point in shadow to the corresponding point on the object. This constraint has been previously used to analyze photos in which corresponding points can be explicitly estimated, but such cases are uncommon in practice. By partially constraining this correspondence to a range of possible points, we developed methods that can detect physical inconsistency when shadows are ambiguously shaped, and under both point and area lights.

In Chapter 3, we developed a method to determine the mutual physical consistency of shading and shadows in photos illuminated by point light sources. To do so, we exploited the well known geometric constraint that the surface normals along the occluding contours of objects can be estimated from a photo. This constraint has been previously used to estimate 2-D light direction under the assumption of orthographic projection. By partially constraining the 3-D light direction with the occluding surface normals, we developed a method that can detect physically inconsistent combinations of shading and shadows in photos captured under linear perspective, and with unknown camera parameters.

In Chapter 4, we reformulated a variety of prior methods in terms of partial constraints,

and showed that they can generalized to analyze photos for which estimation is not possible. Specifically we reformulated methods based upon the camera center of projection, optical chromatic aberrations, and specular reflections. Partial constraints can detect inconsistent centers of projection when only two mutually-orthogonal vanishing points are present (rather than three); inconsistent chromatic aberration when only a few distorted edges are present; and inconsistent specular reflections when the geometry of the reflecting surface is only partially known and imaged under unknown camera parameters.

The limitations of the partial constraint approach can be seen by noting that a particular property may be highly under-constrained in a photo, and this allows many plausible forgeries. For example, a photo with only a few highly ambiguous shadows may suggest many plausible light positions, and shadows can be convincingly falsified if they agree with any such position. This limitation corresponds to the intuition that a photo (or region) may have many plausible interpretations. Partial constraint methods model this space of interpretations, and forgery is detected when a photo specifies conflicting interpretations. This generalizes the approach taken by prior forensic methods. For these the intuition is that a photo (or region) must have only one plausible interpretation, and forgery can be detected when this interpretation is atypical in some way.

Forensic methods based upon partial constraints can be viewed as complimentary to problems in computer vision. An important goal in computer vision is to precisely determine property values for a particular photo. In forensics, however, these property values are somewhat immaterial, as we wish only to determine if a plausible value exists. This relaxed problem is often tractable when the estimation problem is not.

The idea of specifying partial constraints might be used to find new forensic problems or questions. For example, do there exist other properties that have not been considered in the forensics literature because their estimation is known to be difficult or impossible? Partial information about such properties may nonetheless evidence forgery. Similarly, simplified models are often desirable when following an estimation approach (e.g., light direction rather than lighting environments), but such models may discard useful information. Can more complex models of previously-considered properties be used? An example of this latter question is illustrated in Chapter 4, in which some chromatic aberration information was discarded (the expansion factor α of the aberration).

Open forensic problems might also be found by considering whether relaxing the estimation requirement would be beneficial for methods in which estimation is typically possible, but difficult. One example is the case of detecting inconsistent image noise (i.e., photo-response non-uniformity noise). Correlation is estimated in this context [35], but this requires the assumption that forgery occupies a contiguous region of some minimum size. Can partial constraints be used to relax this block-estimation requirement? For other properties, it may be fruitful but difficult to relax the estimation requirement. For example, methods have been developed to detect periodic pixel correlations (e.g., those introduced by resizing

a photo). These methods first estimate a power spectrum, and then detect peaks. Such spectra however contain much information that is ultimately discarded when deciding if a photo is falsified. Is it possible to partially constrain spectra and achieve better detection in difficult cases? Similarly, estimation of camera response functions is difficult, but possible. Would partial constraints achieve more reliable detection when, for example, regions of a photo specify partial information about different portions of the response function? Lastly, region correspondence properties (as used to detect copy-paste forgery) can be difficult to estimate — for example when the copied region has been warped. Might such forgeries be easier to detect if the estimation requirement were relaxed?

By expanding forensic analyses to include photos for which explicit estimation is difficult or impossible, it is our hope that new forensic methods can be designed to detect forgery in a broader range of photos than is currently possible. Although no method can make it impossible to craft a convincing forgery, forensic methods can however require that photographs tell a digitally and physically consistent story by detecting inconsistencies in the information that is present.

Appendix A

Shading

A.1 Line and half-plane constraints

Line constraint

A line segment constraint is defined in terms of a point $\tilde{\boldsymbol{p}}$ and direction $\boldsymbol{d}$:

$$\tilde{\boldsymbol{p}} = \mathbf{K}\mathbf{w}_4 \tag{A.1}$$

$$\boldsymbol{d} = \text{sign}\,(\text{s}_{0_z})(\boldsymbol{p} - \boldsymbol{s}_0) \quad . \tag{A.2}$$

where $\mathbf{K}$ is the intrinsic camera matrix, $\mathbf{w}_4$ is the vector formed by the first three components of the singular vector $\mathbf{v}_4$ from Equation (3.5), and $\boldsymbol{s}_0$ is the least-squares estimate of the 3-D light direction as described in Section 3.1.1.

As can be seen in Figure 3.2(e), the line segment corresponding to the forward constraint (illustrated in red) is simply the line that connects $\boldsymbol{p}$ and $\boldsymbol{s}_0$. The multiplication of the direction $\boldsymbol{d}$ by $\text{sign}\,(\text{s}_{0_z})$ distinguishes between the forward (red) and backward (blue) constraint. That is, when the estimated light direction, $\mathbf{s}_0$, is in front of the camera, $\text{s}_{0_z} < 0$ (by convention) and $\boldsymbol{d}$ is defined as the line segment connecting $\tilde{\boldsymbol{p}}$ to $\boldsymbol{s}_0$. When the estimated light direction is behind the camera, then $\text{s}_{0_z} > 0$ and $\boldsymbol{d}$ is defined as the line segment connecting $\boldsymbol{s}_0$ to $\tilde{\boldsymbol{p}}$.

Half-plane constraint

A half-plane constraint is defined in terms of a point $\tilde{\boldsymbol{p}}$ and direction $\boldsymbol{d}$:

$$\tilde{\boldsymbol{p}} = \mathbf{K}\mathbf{w}_4 \tag{A.3}$$

$$\boldsymbol{d} = \text{sign}\,(\tilde{\boldsymbol{h}} \cdot \tilde{\mathbf{s}}_0) \begin{pmatrix} \tilde{h}_1 & \tilde{h}_2 \end{pmatrix}^T \quad , \tag{A.4}$$

where $\mathbf{K}$ is the intrinsic camera matrix, $\mathbf{w}_4$ is the vector formed by the first three components of the singular vector $\mathbf{v}_4$ from Equation (3.5), $\tilde{\boldsymbol{h}} = \mathbf{K}\mathbf{w}_4 \times \mathbf{K}\mathbf{w}_3$ is the homogeneous line

that connects the images of singular vectors $\mathbf{w}_3$ and $\mathbf{w}_4$, and $\tilde{h}_1$ and $\tilde{h}_2$ are the first and second components of $\tilde{\boldsymbol{h}}$.

As shown in Figure 3.2(f), the line $\tilde{\boldsymbol{h}}$ defines the boundary between the forward (red) and reverse (blue) half-plane constraints. These correspond to the case when the camera is in front of and behind the camera, respectively.

The line $\tilde{\boldsymbol{h}}$ can therefore be used to compute the half-plane direction, $\boldsymbol{d}$. Specifically, $\boldsymbol{d}$ is constructed from the first two components of the line, $\tilde{h}_1$ and $\tilde{h}_2$, which specify a vector that is orthogonal to the half-plane boundary. Multiplication by $\text{sign}\,(\tilde{\boldsymbol{h}} \cdot \tilde{\boldsymbol{s}}_0)$ orients the direction $\boldsymbol{d}$ to identify the side of the half-plane into which light sources in front of the camera project.

To derive the factor $\text{sign}\,(\tilde{\boldsymbol{h}} \cdot \tilde{\boldsymbol{s}}_0)$, note that a light $\boldsymbol{s}_0$ lies on the side ψ of the boundary:

$$\begin{aligned} \psi &= \text{sign}\left(\tilde{\boldsymbol{h}}_{1,2} \cdot \boldsymbol{s}_0 + \tilde{h_3}\right) && \text{(A.5)} \\ &= \text{sign}\left(\tilde{\boldsymbol{h}} \cdot \frac{\tilde{\boldsymbol{s}}_0}{-\text{s}_{0_z}}\right) && \text{(A.6)} \\ &= -\text{sign}\,(\text{s}_{0_z})\text{sign}\left(\tilde{\boldsymbol{h}} \cdot \tilde{\boldsymbol{s}}_0\right) \quad , && \text{(A.7)} \end{aligned}$$

where $\tilde{\boldsymbol{h}}_{1,2}$ is a vector of the first two components of $\tilde{\boldsymbol{h}}$, and the division by $-\text{s}_{0_z}$ normalizes the homogeneous vector $\tilde{\boldsymbol{s}}_0$.

When the light is in front of the camera, it projects to side $\psi = \text{sign}\,(\tilde{\boldsymbol{h}} \cdot \tilde{\boldsymbol{s}}_0)$ because $\text{s}_{0_z} < 0$. If these forward lights project to side $\psi = -1$, then $\text{sign}\,(\tilde{\boldsymbol{h}} \cdot \tilde{\boldsymbol{s}}_0) = -1$ and we define $\boldsymbol{d}$ to point to the opposite side of the half-plane: $\boldsymbol{d} = \text{sign}\,(\tilde{\boldsymbol{h}} \cdot \tilde{\boldsymbol{s}}_0)\tilde{\boldsymbol{h}}_{1,2}$.

Similarly, when the light is behind the camera it projects to side $\psi = -\text{sign}\,(\tilde{\boldsymbol{h}} \cdot \tilde{\boldsymbol{s}}_0)$ because $\text{s}_{0_z} > 0$. If lights behind the camera project to side $\psi = 1$, then $\text{sign}\,(\tilde{\boldsymbol{h}} \cdot \tilde{\boldsymbol{s}}_0) = -1$ and we define $\boldsymbol{d}$ to point to the opposite side of the half-plane: $\boldsymbol{d} = \text{sign}\,(\tilde{\boldsymbol{h}} \cdot \tilde{\boldsymbol{s}}_0)\tilde{\boldsymbol{h}}_{1,2}$.

A.2 Estimating uncertainty functions

The reliability of the estimated line and half-plane constraints described, Section 3.1.2, are impacted by image noise, luminance non-linearities, resolution, and violations of the assumed imaging model. The functions $\Phi_{\text{L}}(\kappa_{\text{L}}, \theta, \rho)$ and $\Phi_{\text{H}}(\kappa_{\text{H}}, \theta, \rho)$, described in Section 3.1.5, model the uncertainty that results from these factors. The specific form of these uncertainty functions is determined through a set of large-scale simulations, as described below.

The uncertainty functions were estimated by simulating images of occluding contours. Specifically, 200 random blobby shapes were generated and imaged under 50 random camera configurations, yielding a total of $10,000$ distinct occluding contours. The focal lengths of the cameras were drawn uniformly at random in the range 18 mm to 200 mm. Each blobby shape was placed randomly in the virtual scene so as to occupy approximately 235×235 pixels in the rendered image.

A total of 3 billion images of the occluding contours were created under differing lighting conditions. Contours were shaded by specifying a random surface albedo, ν in $[0.1,\ 0.5]$. Light directions, $\mathbf{s}$, were chosen randomly from the unit sphere, excluding directions more than 45° behind the contour. The contour was shaded according to the Lambertian reflectance equation, $r = \nu(\mathbf{s} \cdot \mathbf{n})$. To this primary illumination, a secondary light was added to simulate inter-reflections and other effects. Secondary lights were drawn randomly from the unit sphere, and their peak intensity on the occluding contour was, on average, 15% of the peak primary radiance (standard deviation 5%). An ambient illumination was also added, drawn from a normal distribution with mean 0 and standard deviation 0.1 (negative values were excluded). Lastly, the final rendered image was subjected to 1% additive Gaussian noise, gamma correction $r^{1/\gamma}$ with γ drawn uniformly in $[1.2,\ 1.8]$, and 8-bit quantization. See Appendix A.3 for details.

The 3 billion images afforded a dense sampling of the wide range of line and half-plane constraints that result when image and lighting conditions vary, and when portions of the contour are obscured in an image. The illuminated portion of the contour in each image was divided into $10 - 20$ segments, a random subset of which were used to build line and half-plane constraints. The 3-D normals on the contour were estimated, Equation (3.30), by selecting a random focal length, f, within δ_f log-units of the true focal length, and a random image center, $\boldsymbol{c}$, within $\delta_{\boldsymbol{c}}$ units of the true image center. (The values of δ_f and $\delta_{\boldsymbol{c}}$ are discussed below.) The true image center was selected uniformly at random within a circle with diameter 12% of the image height. The occluding normals were used to construct the linear system, and thus the line constraint, $(\boldsymbol{p}_{\text{L}},\ \boldsymbol{d}_{\text{L}})$, and half-plane constraint, $(\boldsymbol{p}_{\text{H}},\ \boldsymbol{d}_{\text{H}})$. The angle between the constraint directions $\boldsymbol{d}$ and the vector from $\boldsymbol{p}$ to the light source was then computed, up to an inversion of $\boldsymbol{d}$ for lights behind the camera. This produced 3 billion samples of these angles, ϕ_{L} and ϕ_{H}, from which the functions, Φ_{L} and Φ_{H}, were estimated.

The functions, $\Phi_{\text{L}}(\kappa_{\text{L}}, \theta, \rho)$ and $\Phi_{\text{H}}(\kappa_{\text{H}}, \theta, \rho)$, were estimated by quantizing the function domain and computing a confidence on the values ϕ_{L} and ϕ_{H} associated with each quantized location. The condition, $\kappa_{\text{L}} = \lambda_3/\lambda_1$ and $\kappa_{\text{H}} = \lambda_2/\lambda_1$, was quantized, in the log domain, into 24 levels between 0.001 and 0.7 (the minimum and maximum practical values that were encountered). The number of surface normal directions, θ, was measured by computing the angle of each 2-D image normal, constructing a histogram of these angles, and computing the number of degrees comprised by the non-empty bins. This number of directions, θ, was quantized into 36 levels (10° increments between 0° and 360°). The R-value, ρ, of the least squares solution, Equation (3.4), was quantized into 30 levels between 0.7 and 1. Note that these particular quantization ranges and levels are not essential, and were chosen to give reasonable resolution and intuitive quantization steps. The value of the functions, Φ_{L} and Φ_{H}, at each quantized position specified the angle below which 99.9% of the samples, ϕ_{L} and ϕ_{H}, fell.

Camera precision	$1/6^{\text{th}}$	$1/3^{\text{rd}}$	1
Mean (std-dev), Φ_{L}	33° (18°)	33° (18°)	58° (30°)
Mean (std-dev), Φ_{H}	96° (5°)	96° (5°)	96° (5°)

Figure A.1: The effect of varying precision in the focal length, f, and image center, $\boldsymbol{c}$, on the average value of the wedge functions, Φ_{L} and Φ_{H}. Precision is measured as a fraction of the plausible range of the log focal length, 18mm–200mm, and image center, $\pm 12\%$ of the image height. Estimates of f and $\boldsymbol{c}$ that are more precise than $1/3^{\text{rd}}$ do not affect average constraint width.

Recall that the functions Φ_{L} and Φ_{H} are estimated by assuming that the camera focal length, f and image center $\boldsymbol{c}$ are known within δ_f and $\delta_{\boldsymbol{c}}$ units of their true values. The effect of these precisions on the functions was tested by estimating the functions at varying values of δ_f and $\delta_{\boldsymbol{c}}$. Three precisions were selected, corresponding to fractions of the plausible range of the log focal length and of the image center. (Recall that plausible focal lengths were in [18 200]mm and plausible image centers were within 12% of the image height.) The precisions were $1/6^{\text{th}}$, $1/3^{\text{rd}}$, and 1 (the entire range). The focal precisions were $\delta_f = \pm(0.087,\ 0.174,\ 0.522)$ log focal units. The image center precisions were $\delta_{\boldsymbol{c}} = \pm(2\%,\ 4\%,\ 12\%)$ of the image height.

The effect of the precision of the focal length and image center was tested by computing the mean angle in the functions, Figure A.1. When the focal length and image center are known within $1/6^{\text{th}}$ or $1/3^{\text{rd}}$ of the plausible range, this mean half-angle was the same: 33° for Φ_{L} and 96° for Φ_{H}. The mean half-angle of Φ_{L} increases significantly for a fraction of 1, suggesting that estimates of the focal length within $1/3^{\text{rd}}$ of the plausible range are sufficient, ($\delta_f = \pm 0.174$ and $\delta_{\boldsymbol{c}} = \pm 4\%$).

A.3 Shading of occluding contours

The creation of 3 billion images of random occluding contours was simulated under varying lighting conditions. A primary light source $\mathbf{s}$, secondary light source $\mathbf{s}_2$, and an ambient illumination a were used to shade the occluding contours of the objects shown in Figure 3.6. The radiance contributed by the primary and secondary light sources were given by:

$$r_1 = \nu(\mathbf{s} \cdot \mathbf{n}) \tag{A.8}$$

$$r_2 = \nu(\mathbf{s}_2 \cdot \mathbf{n}) \ , \tag{A.9}$$

where $\mathbf{n}$ is the surface normal, ν is the surface albedo drawn uniformly and at random from the range $[0.1,\ 0.5]$, and where any radiances less than zero are clipped to zero. These light

sources were combined as follows:

$$r_c \;=\; (1-\alpha)r_1 + \alpha r_2 \quad , \tag{A.10}$$

where the contribution, $\alpha \in [0,\ 1]$, of the secondary light source accounted for only a fraction of the observed image intensity specified as:

$$\frac{\alpha \max(r_2)}{(1-\alpha)\max(r_1)} \;=\; \tau \quad , \tag{A.11}$$

with the max computed along the portion of the contour illuminated by the primary light source, and the fraction τ drawn from a normal distribution with mean 0.15 and standard deviation 0.05.

The ambient illumination, a, was added to this combined radiance:

$$r \;=\; \frac{r_c + a}{0.5 + a} \quad , \tag{A.12}$$

where 0.5 is the maximum possible radiance due to the primary and secondary light sources and albedo, ν. This divisive normalization simulates a camera exposure that accommodates the brightest possible occluding contour. The final radiance, r, is therefore bounded in the range $[0,\ 1]$.

The image of the occluding contour was computed from these shaded contours by adding 1% Gaussian noise η, random gamma of the form $r^{1/\gamma}$ with γ drawn uniformly and at random in the range $[1.2,\ 1.8]$, and then subjected to quantization:

$$\hat{r} \;=\; Q\big((Q(r+\eta))^{1/\gamma}\big) \quad , \tag{A.13}$$

where the function $Q(\cdot)$ quantizes the signal into 256 levels.

A.4 Occluding surface normals

Occluding contours typically form at the boundary of objects in an image. The camera ray to points along these contours is tangent to the surface and orthogonal to the surface normal. As such, surface normals along occluding contours comprise only two degrees of freedom. Under orthographic projection, the z-component of these occluding surface normals is zero, and the x, y components are given by the image normal [105, 185]. We however need to specify these normals under a model of perspective projection.

Under perspective projection, camera rays to points along an occluding contour span a hemisphere of directions, but are projected onto a plane. This planar projection distorts the shape of the occluding contour and obscures the 3-D orientation of the surface normals. These properties can be recovered by mapping the planar image onto the portion of the hemisphere that it subtends.

Specifically, we map the planar image I onto a spherical image, $\hat{I}$, of radius r:

$$\hat{I}(\theta, \phi, r) \;=\; I\big(\boldsymbol{x}(\theta, \phi, r)\big) \;, \tag{A.14}$$

where (θ, ϕ, r) is the spherical coordinate of a point in the spherical image. The corresponding point in the planar image is $\boldsymbol{x}(\theta, \phi, r)$, and is the projection of the 3-D spherical point $\mathbf{x}(\theta, \phi, r)$ into I,

$$\boldsymbol{x}(\theta, \phi, r) \;=\; \frac{-f}{\mathrm{x}_z(\theta, \phi, r)} \begin{bmatrix} \mathrm{x}_x(\theta, \phi, r) \\ \mathrm{x}_y(\theta, \phi, r) \end{bmatrix} + \boldsymbol{c} \;, \tag{A.15}$$

under a camera with focal length f and image center $\boldsymbol{c}$. The 3-D point $\mathbf{x}(\theta, \phi, r)$ — referenced above by subscripts x, y, z — is the cartesian equivalent of the point on the sphere:

$$\mathbf{x}(\theta, \phi, r) \;=\; \begin{bmatrix} r\cos\theta\sin\phi \\ r\sin\theta\sin\phi \\ r\cos\phi \end{bmatrix} \;. \tag{A.16}$$

Points $\boldsymbol{x}$ are therefore

$$\boldsymbol{x}(\theta, \phi, r) \;=\; \frac{-f}{r\cos\phi} \begin{bmatrix} r\cos\theta\sin\phi \\ r\sin\theta\sin\phi \end{bmatrix} + \boldsymbol{c} \tag{A.17}$$

$$\;=\; -f \begin{bmatrix} \cos\theta\tan\phi \\ \sin\theta\tan\phi \end{bmatrix} + \boldsymbol{c} \;\;, \tag{A.18}$$

where the radius of the spherical image, r, does not affect the projection of $\mathbf{x}(\theta, \phi, r)$ to $\boldsymbol{x}(\theta, \phi, r)$.

Given the spherical image $\hat{I}$, computing the 3-D surface normal of an occluding contour under perspective projection is analogous to the orthographic case: the occluding normal is the direction in this spherical image that is orthogonal to the image of the contour. This direction can be derived from the gradient of $\hat{I}(\theta, \phi, r)$ (without sacrificing generality). The spherical gradient is

$$\nabla\hat{I}(\theta, \phi, r) \;=\; \frac{1}{r\sin\phi}\frac{\partial\hat{I}}{\partial\theta}\hat{\boldsymbol{\theta}} \;+\; \frac{1}{r}\frac{\partial\hat{I}}{\partial\phi}\hat{\boldsymbol{\phi}} \;+\; \frac{\partial\hat{I}}{\partial r}\hat{\boldsymbol{r}} \tag{A.19}$$

$$\;=\; \begin{bmatrix} \dfrac{\partial\hat{I}}{\partial\theta} & \dfrac{\partial\hat{I}}{\partial\phi} & \dfrac{\partial\hat{I}}{\partial r} \end{bmatrix} \begin{bmatrix} \frac{1}{r\sin\phi} & 0 & 0 \\ 0 & \frac{1}{r} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\theta}}^{\,T} \\ \hat{\boldsymbol{\phi}}^{\,T} \\ \hat{\boldsymbol{r}}^{T} \end{bmatrix}, \tag{A.20}$$

where the diagonal matrix normalizes the scale of the partial derivatives along the spherical basis directions $\hat{\boldsymbol{r}}$, $\hat{\boldsymbol{\theta}}$, and $\hat{\boldsymbol{\phi}}$. The partials are given by the general chain rule:

$$\begin{bmatrix} \frac{\partial \hat{I}}{\partial \theta} & \frac{\partial \hat{I}}{\partial \phi} & \frac{\partial \hat{I}}{\partial r} \end{bmatrix} = \begin{bmatrix} \frac{\partial I}{\partial x_x} & \frac{\partial I}{\partial x_y} \end{bmatrix} \begin{bmatrix} \frac{\partial x_x}{\partial \theta} & \frac{\partial x_x}{\partial \phi} & \frac{\partial x_x}{\partial r} \\ \frac{\partial x_y}{\partial \theta} & \frac{\partial x_y}{\partial \phi} & \frac{\partial x_y}{\partial r} \end{bmatrix} \tag{A.21}$$

$$= \nabla I \begin{bmatrix} f \sin\theta \tan\phi & -f \cos\theta \sec^2\phi & 0 \\ -f \cos\theta \tan\phi & -f \sin\theta \sec^2\phi & 0 \end{bmatrix}. \tag{A.22}$$

Combining Equation (A.20) and Equation (A.22) and reducing gives

$$\nabla \hat{I}(\theta, \phi, r) = \nabla I \begin{bmatrix} \sin\theta & \cos\theta \\ -\cos\theta & \sin\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & r/f \end{bmatrix} \begin{bmatrix} -\hat{\boldsymbol{\theta}}^T \\ -\hat{\boldsymbol{\phi}}^T \end{bmatrix}, \tag{A.23}$$

where the radial component $\hat{\boldsymbol{r}}$ does not appear because the radius of the spherical image does not affect the image content. The negation of the bases accounts for the negation during projection, Equation (A.18). See Appendix A.5 for details of this reduction.

Three transformations are therefore required to map the 2-D image gradient, ∇I, into a 3-D vector that is parallel to the 3-D occluding surface normal. First the 2-D gradient is rotated, within the image, into a local 2-D coordinate system whose axes are aligned with x, y components of the 3-D basis vectors. The $\hat{\boldsymbol{\theta}}$ basis vector is parallel to the image plane (its z-component is zero), and it is oriented in an axial direction around the image center. The x, y components of the $\hat{\boldsymbol{\phi}}$ basis vector are oriented in a radial direction away from the image center, and the z-component is nonzero. This radial component of the image gradient is next scaled by r/f, which corrects for the geometric distortion in the image of the local region. This distortion stretches the image when it is projected onto the image plane, which shortens the radial component of the 3-D gradient by f/r, the cosine of the angle between the image plane and the plane spanned by the spherical basis vectors. Lastly, the scaled 2-D image gradient is mapped into 3-D as a weighted combination of the 3-D basis vectors, $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\phi}}$.

Equation (A.23) can be further simplified to eliminate the spherical coordinates

$$\nabla \hat{I}(\theta, \phi, r) = \nabla I \begin{bmatrix} 1 & 0 & (x_x - c_x)/f \\ 0 & 1 & (x_y - c_y)/f \end{bmatrix} \tag{A.24}$$

$$= \begin{bmatrix} \nabla I & \frac{1}{f} \nabla I (\boldsymbol{x} - \boldsymbol{c}) \end{bmatrix}, \tag{A.25}$$

and the resulting expression applies to any measure of the 2-D normal along an occluding

contour in an image. The 3-D surface normal $\mathbf{n}$ of an occluding contour under perspective projection is

$$\mathbf{n} = \begin{bmatrix} n_x \\ n_y \\ \frac{1}{f}\,\boldsymbol{n}\cdot(\boldsymbol{x}-\boldsymbol{c}) \end{bmatrix}, \tag{A.26}$$

where $\boldsymbol{n} = \begin{pmatrix} n_x & n_y \end{pmatrix}^T$ is the 2-D normal of the occluding contour in the image, $\boldsymbol{x}$ is the position of the 2-D normal in the image, f is the focal length, and $\boldsymbol{c}$ is the image center.

Under orthographic projection, $f \to \infty$ and $\mathbf{n} \to \begin{pmatrix} n_x & n_y & 0 \end{pmatrix}^T$, as expected. Similarly, the z-component of $\mathbf{n}$ vanishes at the image center, $\boldsymbol{x} - \boldsymbol{c} = 0$. The direction of the x, y components of $\mathbf{n}$ are unchanged by projection. Note that this expression for the normal does not have unit length — even if $\boldsymbol{n}$ does — because the planar image projects onto the sphere with spatially-varying density, and this modulates the gradient magnitude.

A.5 Occluding surface normals (algebraic reduction)

Here we perform the algebraic reduction that is required to derive the occluding surface normal, as summarized in Appendix A.4. Starting from Equation (A.19), the gradient of the spherical image $\hat{I}$ is given by

$$\nabla\hat{I}(\theta,\phi,r) = \frac{1}{r\sin\phi}\frac{\partial\hat{I}}{\partial\theta}\hat{\boldsymbol{\theta}} + \frac{1}{r}\frac{\partial\hat{I}}{\partial\phi}\hat{\boldsymbol{\phi}} + \frac{\partial\hat{I}}{\partial r}\hat{\boldsymbol{r}} \tag{A.27}$$

$$= \begin{bmatrix} \frac{\partial\hat{I}}{\partial\theta} & \frac{\partial\hat{I}}{\partial\phi} & \frac{\partial\hat{I}}{\partial r} \end{bmatrix} \begin{bmatrix} \frac{1}{r\sin\phi} & 0 & 0 \\ 0 & \frac{1}{r} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\theta}}^T \\ \hat{\boldsymbol{\phi}}^T \\ \hat{\boldsymbol{r}}^T \end{bmatrix} \tag{A.28}$$

where $\hat{\boldsymbol{r}}$, $\hat{\boldsymbol{\theta}}$, and $\hat{\boldsymbol{\phi}}$ are the spherical basis vectors (defined later). The partials are given by the general chain rule:

$$\begin{bmatrix} \frac{\partial\hat{I}}{\partial\theta} & \frac{\partial\hat{I}}{\partial\phi} & \frac{\partial\hat{I}}{\partial r} \end{bmatrix} = \begin{bmatrix} \frac{\partial I}{\partial x_x} & \frac{\partial I}{\partial x_y} \end{bmatrix} \begin{bmatrix} \frac{\partial x_x}{\partial\theta} & \frac{\partial x_x}{\partial\phi} & \frac{\partial x_x}{\partial r} \\ \frac{\partial x_y}{\partial\theta} & \frac{\partial x_y}{\partial\phi} & \frac{\partial x_y}{\partial r} \end{bmatrix} \tag{A.29}$$

$$= \nabla I(x,y) \begin{bmatrix} f\sin\theta\tan\phi & -f\cos\theta\sec^2\phi & 0 \\ -f\cos\theta\tan\phi & -f\sin\theta\sec^2\phi & 0 \end{bmatrix}. \tag{A.30}$$

Combining these expressions gives

$$\nabla \hat{I}(\theta,\phi,r) = \nabla I \begin{bmatrix} f\sin\theta\tan\phi & -f\cos\theta\sec^2\phi & 0 \\ -f\cos\theta\tan\phi & -f\sin\theta\sec^2\phi & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{r\sin\phi} & 0 & 0 \\ 0 & \frac{1}{r} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\theta}}^T \\ \hat{\boldsymbol{\phi}}^T \\ \hat{\boldsymbol{r}}^T \end{bmatrix} \quad (A.31)$$

$$= \nabla I \begin{bmatrix} f\sin\theta\tan\phi \Big/ r\sin\phi & -f\cos\theta\sec^2\phi/r \\ -f\cos\theta\tan\phi \Big/ r\sin\phi & -f\sin\theta\sec^2\phi/r \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\theta}}^T \\ \hat{\boldsymbol{\phi}}^T \end{bmatrix}. \quad (A.32)$$

This expression can be reduced by substituting for the spherical coordinates

$$\theta = \arctan 2\Big[x_y - c_y,\ \ x_x - c_x\Big] \quad (A.33)$$

$$\phi = \arccos\big[-f/r\big] \quad (A.34)$$

$$r = \sqrt{(x_x - c_x)^2 + (x_y - c_y)^2 + f^2}\ , \quad (A.35)$$

which are centered at point $(c_x,\ \ c_y,\ \ 0)^T$ in cartesian space (f units away from the center of the planar image). Simplifying Equation (A.32) gives

$$\nabla \hat{I}(\theta,\phi,r) = \nabla I \begin{bmatrix} -\sin\theta & -r\cos\theta/f \\ \cos\theta & -r\sin\theta/f \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\theta}}^T \\ \hat{\boldsymbol{\phi}}^T \end{bmatrix}. \quad (A.36)$$

$$= \nabla I \begin{bmatrix} \sin\theta & \cos\theta \\ -\cos\theta & \sin\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & r/f \end{bmatrix} \begin{bmatrix} -\hat{\boldsymbol{\theta}}^T \\ -\hat{\boldsymbol{\phi}}^T \end{bmatrix}\ , \quad (A.37)$$

which is the first expression of the gradient in Appendix A.4 — Equation (A.23). This expression can be further reduced to eliminate the spherical coordinate representation by substituting for the basis vectors $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\phi}}$. The basis vectors are

$$\hat{\boldsymbol{\theta}}^T = \Big[-\sin\theta \quad \cos\theta \quad 0\Big] \quad (A.38)$$

$$\hat{\boldsymbol{\phi}}^T = \Big[\cos\theta\cos\phi \quad \sin\theta\cos\phi \quad -\sin\phi\Big]\ , \quad (A.39)$$

and substitution gives

$$\nabla \hat{I}(\theta,\phi,r) = \nabla I \begin{bmatrix} \sin\theta & \cos\theta \\ -\cos\theta & \sin\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & r/f \end{bmatrix} \begin{bmatrix} \sin\theta & -\cos\theta & 0 \\ -\cos\theta\cos\phi & -\sin\theta\cos\phi & \sin\phi \end{bmatrix} \quad (A.40)$$

$$= \nabla I \begin{bmatrix} \sin\theta & \cos\theta \\ -\cos\theta & \sin\theta \end{bmatrix} \begin{bmatrix} \sin\theta & -\cos\theta & 0 \\ \cos\theta & \sin\theta & r\sin\phi/f \end{bmatrix} \quad (A.41)$$

$$= \nabla I \begin{bmatrix} 1 & 0 & r\cos\theta\sin\phi/f \\ 0 & 1 & r\sin\theta\sin\phi/f \end{bmatrix} , \quad (A.42)$$

where the spherical coordinate expressions have again been used in the reduction. The remaining trigonometric functions can be eliminated by noting that

$$\cos\theta = \cos\Big[\arctan2\big(x_y - c_y,\ \ x_x - c_x\big)\Big] = \frac{x_x - c_x}{\|\boldsymbol{x} - \boldsymbol{c}\|} \quad (A.43)$$

$$\sin\theta = \sin\Big[\arctan2\big(x_y - c_y,\ \ x_x - c_x\big)\Big] = \frac{x_y - c_y}{\|\boldsymbol{x} - \boldsymbol{c}\|} \quad (A.44)$$

$$\sin\phi = \sin\Big[\arccos\big(-f/r\big)\Big] = \sqrt{1 - f^2/r^2}\ , \quad (A.45)$$

where the first two result from the definition of the tangent, sin, and cosine, and the last from the pythagorean theorem. Substituting and reducing the upper term in Equation (A.42)

$$\frac{r}{f}\cos\theta\sin\phi = \frac{r}{f}\frac{x_x - c_x}{\sqrt{(x_x - c_x)^2 + (x_y - c_y)^2}}\sqrt{1 - f^2/r^2} \quad (A.46)$$

$$= \frac{1}{f}\frac{x_x - c_x}{\sqrt{(x_x - c_x)^2 + (x_y - c_y)^2}}\sqrt{r^2 - f^2} \quad (A.47)$$

$$= \frac{1}{f}\frac{x_x - c_x}{\sqrt{(x_x - c_x)^2 + (x_y - c_y)^2}}\sqrt{(x_x - c_x)^2 + (x_y - c_y)^2 + f^2 - f^2} \quad (A.48)$$

$$= \frac{x_x - c_x}{f}\ . \quad (A.49)$$

Similar algebra applies to the lower term in Equation (A.42), so

$$\frac{r}{f}\sin\theta\sin\phi = \frac{x_y - c_y}{f}\ . \quad (A.50)$$

Equations (A.49) and (A.50) eliminate the trigonometric functions in Equation (A.42):

$$\nabla \hat{I}(\theta, \phi, r) = \nabla I \begin{bmatrix} 1 & 0 & (x_x - c_x)/f \\ 0 & 1 & (x_y - c_y)/f \end{bmatrix} \tag{A.51}$$

$$= \begin{bmatrix} \nabla I & \frac{1}{f} \nabla I \cdot \big(\boldsymbol{x} - \boldsymbol{c}\big) \end{bmatrix}, \tag{A.52}$$

and give a concise relationship between the 2-D image gradient ∇I and the 3-D gradient in the spherical image. This result applies to any measure of the 2-D normal $\boldsymbol{n}$ along an occluding contour. In general,

$$\mathbf{n} = \begin{bmatrix} n_x \\ n_y \\ \frac{1}{f}\, \boldsymbol{n} \cdot (\boldsymbol{x} - \boldsymbol{c}) \end{bmatrix}, \tag{A.53}$$

where $\mathbf{n}$ is the 3-D normal on the occluding contour, $\boldsymbol{x}$ is the position of the 2-D normal in the image, f is the focal length, and $\boldsymbol{c}$ is the image center.

Bibliography

[1] P. Alvarez. Using extended file information (EXIF) file headers in digital evidence analysis. *International Journal of Digital Evidence*, 2(3):1–5, 2004.

[2] I. Amerini, L. Ballan, R. Caldelli, A. Del Bimbo, and G. Serra. A SIFT-based forensic method for copy–move attack detection and transformation recovery. *IEEE Transactions on Information Forensics and Security*, 6(3):1099–1110, 2011.

[3] I. Amerini, L. Ballan, R. Caldelli, A. Del Bimbo, and G. Serra. Geometric tampering estimation by means of a SIFT-based forensic analysis. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 1702–1705, 2010.

[4] E. Ardizzone, A. Bruno, and G. Mazzola. Detecting multiple copies in tampered images. In *International Conference on Image Processing*, 2010.

[5] I. Avcibas, S. Bayram, N. Memon, M. Ramkumar, and B. Sankur. A classifier design for detecting image manipulations. In *IEEE International Conference on Image Processing*, pages 2645–2648, 2004.

[6] M. Barni, A. Costanzo, and L. Sabatini. Identification of cut & paste tampering by means of double-JPEG detection and image segmentation. In *International Symposium on Circuits and Systems*, 2010.

[7] M. K. Bashar, K. Noda, N. Ohnishi, H. Kudo, T. Matsumoto, and Y. Takeuchi. Wavelet-based multiresolution features for detecting duplications in images. In *Machine Vision Application*, pages 264–267, 2007.

[8] M. K. Bashar, K. Noda, N. Ohnishi, and K. Mori. Exploring duplicated regions in natural images. *IEEE Transactions on Image Processing*, PP(99):1, 2010.

[9] R. Basri and D. Jacobs. Lambertian reflectance and linear subspaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(2):218–233, 2003.

[10] S. Bayram, I. Avcibas, B. Sankur, and N. Memon. Image manipulation detection with binary similarity measures. In *European Signal Processing Conference*, pages 752–755, 2005.

[11] S. Bayram, I. Avcibas, B. Sankur, and N. Memon. Image manipulation detection. *Journal of Electronic Imaging*, 15(4):041102, 2006.

[12] S. Bayram, H. Sencar, and N. Memon. A survey of copy-move forgery detection techniques. In *IEEE Western New York Image Processing Workshop*, 2008.

[13] S. Bayram, H. T. Sencar, and N. Memon. An efficient and robust method for detecting copy-move forgery. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 1053–1056, 2009.

[14] P. Beckmann. *Scattering of Electromagnetic Waves from Rough Surfaces*. Number 0-89006-238-2. Artech House Publishers, New York, NY, 1987.

[15] T. Bianchi and A. Piva. Detection of non-aligned double JPEG compression based on integer periodicity maps. *IEEE Transactions on Information Forensics and Security*, 7(2):842–848, 2012.

[16] T. Bianchi and A. Piva. Image forgery localization via block-grained analysis of JPEG artifacts. *IEEE Transactions on Information Forensics and Security*, 7(3):1003–1017, 2012.

[17] T. Bianchi, A. De Rosa, and A. Piva. Improved DCT coefficient analysis for forgery localization in JPEG images. In *International Conference on Acoustics, Speech, and Signal Processing*, 2011.

[18] X. Bo, W. Junwen, L. Guangjie, and D. Yuewei. Image copy-move forgery detection based on SURF. In *International Conference on Multimedia Information Networking and Security, 2010*, pages 889–892, 2010.

[19] A. Bovik. Streaking in median filtered images. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 35(4):493–503, 1987.

[20] S. Bravo-Solorio and A. K. Nandi. Passive method for detecting duplicated regions affected by reflection, rotation and scaling. In *European Signal Processing Conference*, pages 824–828, 2009.

[21] S. Bravo-Solorio and A. K. Nandi. Automated detection and localisation of duplicated regions affected by reflection, rotation and scaling in image forensics. *Signal Processing*, 91(8):1759–1770, 2011.

[22] S. Bravo-Solorio and A.K. Nandi. Exposing duplicated regions affected by reflection, rotation and scaling. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 1880–1883, 2011.

[23] R. Caldelli, I. Amerini, and A. Novi. An analysis on attacker actions in fingerprint-copy attack in source camera identification. In *IEEE International Workshop on Information Forensics and Security*, pages 1–6, 2011.

[24] G. Cao, Y. Zhao, and R. Ni. Detection of image sharpening based on histogram aberration and ringing artifacts. In *IEEE International Conference on Multimedia and Expo*, pages 1026–1029, 2009.

[25] G. Cao, Y. Zhao, and R. Ni. Forensic estimation of gamma correction in digital images. In *IEEE International Conference on Image Processing*, pages 2097–2100, 2010.

[26] G. Cao, Y. Zhao, and R. Ni. Forensic identification of resampling operators: A semi non-intrusive approach. *Forensic Science International*, 216:29–36, 2012.

[27] G. Cao, Y. Zhao, R. Ni, and A.C. Kot. Unsharp masking sharpening detection via overshoot artifacts analysis. *IEEE Signal Processing Letters*, 18(10):603–606, 2011.

[28] G. Cao, Y. Zhao, R. Ni, and H. Tian. Anti-forensics of contrast enhancement in digital images. In *Proceedings of the 12th ACM workshop on Multimedia and Security*, pages 25–34, New York, NY, USA, 2010. ACM.

[29] G. Cao, Y. Zhao, R. Ni, L. Yu, and H. Tian. Forensic detection of median filtering in digital images. In *IEEE International Conference on Multimedia and Expo*, pages 89–94, 2010.

[30] H. Cao and A. C. Kot. Manipulation detection on image patches using FusionBoost. *IEEE Transactions on Information Forensics and Security*, 7(3):992–1002, 2012.

[31] C. Chen and J. Ni. Median filtering detection using edge based prediction matrix. In YunQing Shi, Hyoung-Joong Kim, and Fernando Perez-Gonzalez, editors, *Digital Forensics and Watermarking*, volume 7128 of *Lecture Notes in Computer Science*, pages 361–375. Springer Berlin Heidelberg, 2012.

[32] C. Chen, J. Ni, R. Huang, and J. Huang. Blind median filtering detection using statistics in difference domain. In *International Workshop on Information Hiding*, 2012.

[33] D. Chen, J. Li, S. Wang, and S. Li. Identifying computer generated and digital camera images using fractional lower order moments. In *IEEE Conference on Industrial Electronics and Applications*, pages 230–235, 2009.

[34] M. Chen, J. Fridrich, M. Goljan, and J. Lukáš. Determining image origin and integrity using sensor noise. *IEEE Transactions on Information Forensics and Security*, 3(1):74–90, 2008.

[35] M. Chen, J. Fridrich, J. Lukáš, and M. Goljan. Imaging sensor noise as digital x-ray for revealing forgeries. In *International Workshop on Information Hiding*, pages 342–358, 2007.

[36] W. Chen, Y. Q. Shi, and W. Su. Image splicing detection using 2-D phase congruency and statistical moments of characteristic function. In Edward J. Delp and Ping Wah Wong, editors, *SPIE Conference on Security, Steganography, and Watermarking of Multimedia Contents*, volume 6505, 2007.

[37] W. Chen, Y. Q. Shi, and G. R. Xuan. Identifying computer graphics using HSV color model and statistical moments of characteristic functions. In *IEEE International Conference on Multimedia and Expo*, pages 1123–1126, 2007.

[38] Y. Chen and C. Hsu. Image tampering detection by blocking periodicity analysis in JPEG compressed images. In *IEEE Workshop on Multimedia Signal Processing*, pages 803–808, 2008.

[39] Y. Chen and C. Hsu. Detecting doubly compressed images based on quantization noise model and image restoration. In *IEEE Workshop on Multimedia Signal Processing*, pages 1–6, 2009.

[40] Y. Chen and C. Hsu. Detecting recompression of JPEG images via periodicity analysis of compression artifacts for tampering detection. *IEEE Transactions on Information Forensics and Security*, 6(2):396–406, 2011.

[41] H. R. Chennamma and L. Rangarajan. Image splicing detection using inherent lens radial distortion. *International Journal of Computer Science Issues*, 7(6):149–158, 2010.

[42] G. Chierchia, S. Parrilli, G. Poggi, C. Sansone, and L. Verdoliva. On the influence of denoising in PRNU based forgery detection. In *Proceedings of the 2nd ACM workshop on Multimedia in forensics, security, and intelligence*, MiFor '10, pages 117–122, New York, NY, USA, 2010. ACM.

[43] G. Chierchia, S. Parrilli, G. Poggi, L. Verdoliva, and C. Sansone. PRNU-based detection of small-size image forgeries. In *International Conference on Digital Signal Processing*, pages 1–6, 2011.

[44] V. Christlein, C. Riess, and E. Angelopoulou. On rotation invariance in copy-move forgery detection. In *IEEE International Workshop on Information Forensics and Security*, pages 1–6, 2010.

[45] V. Christlein, C. Riess, and E. Angelopoulou. A study on features for the detection of copy-move forgeries. In *Information Security Solutions Europe*, 2010.

[46] V. Christlein, C. Riess, J. Jordan, C. Riess, and E. Angelopoulou. An evaluation of popular copy-move forgery detection approaches. *IEEE Transaction on Information Forensics and Security*, 7(6):1841–1854, 2012.

[47] W. Chuang, A. Swaminathan, and M. Wu. Tampering identification using empirical frequency response. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 1517–1520, 2009.

[48] V. Conotter, G. Boato, and H. Farid. Detecting photo manipulation on signs and billboards. In *IEEE International Conference on Image Processing*, pages 1741–1744, 2010.

[49] V. Conotter and L. Cordin. Detecting photographic and computer generated composites. In *SPIE Conference on Media Forensics and Security*, 2011.

[50] N. Dalgaard, C. Mosquera, and F. Perez-Gonzalez. On the role of differentiation for resampling detection. In *IEEE International Conference on Image Processing*, pages 1753–1756, 2010.

[51] V. David and P. Fernando. Prefilter design for forensic resampling estimation. In *Workshop on Information Forensics and Security*, pages 1–6, 2011.

[52] V. David, C. Mosquera, and P. Fernando. Two-dimensional statistical test for the presence of almost cyclostationarity on images. In *International Conference on Image Processing*, pages 1745–1748, 2010.

[53] S. Dehnie, H. T. Sencar, and N. Memon. Digital image forensics for identifying computer generated and digital camera images. In *IEEE International Conference on Image Processing*, pages 2313–2316, 2006.

[54] A. E. Dirik, S. Bayram, H. T. Sencar, and N. Memon. New features to identify computer generated images. In *IEEE International Conference on Image Processing*, volume 4, pages IV–433–IV–436, 2007.

[55] A. E. Dirik and N. Memon. Image tamper detection based on demosaicing artifacts. In *IEEE International Conference on Image Processing*, pages 1509–1512, 2009.

[56] J. Dong, W. Wang, T. Tan, and Y. Q. Shi. Run-length and edge statistics based approach for image splicing detection. In *International Workshop on Digital Watermarking*, volume 5450, pages 76–87, 2009.

[57] B. Dybala, B. Jennings, and D. Letscher. Detecting filtered cloning in digital images. In *Proceedings of the 9th workshop on Multimedia & security*, MM&Sec, pages 43–50, New York, NY, USA, 2007. ACM.

[58] J. Fan, H. Cao, and A.C. Kot. Estimating EXIF parameters based on noise features for image manipulation detection. *IEEE Transactions on Information Forensics and Security*, 8(4):608–618, 2013.

[59] J. Fan, A. C. Kot, H. Cao, and F. Sattar. Modeling the EXIF-image correlation for image manipulation detection. In *IEEE International Conference on International Conference on Image Processing*, pages 1945–1948, 2011.

[60] N. Fan, C. Jin, and Y. Huang. A pixel-based digital photo authentication framework via demosaicking inter-pixel correlation. In *Proceedings of the 11th ACM workshop on Multimedia and security*, MM&Sec, pages 125–130, New York, NY, USA, 2009. ACM.

[61] H. Farid. Detecting digital forgeries using bispectral analysis. Technical Report AIM-1657, AI Lab, Massachusetts Institute of Technology, 1999.

[62] H. Farid. Digital image ballistics from JPEG quantization. Technical Report TR2006-583, Department of Computer Science, Dartmouth College, 2006.

[63] H. Farid. Digital image ballistics from JPEG quantization: A followup study. Technical Report TR2008-638, Department of Computer Science, Dartmouth College, 2008.

[64] H. Farid. Exposing digital forgeries from JPEG ghosts. *IEEE Transactions on Information Forensics and Security*, 1(4):154–160, 2009.

[65] H. Farid. A survey of image forgery detection. *IEEE Signal Processing Magazine*, 2(26):16–25, 2009.

[66] H. Farid and M. J. Bravo. Image forensic analyses that elude the human visual system. In *SPIE Conference on Media Forensics and Security*, 2010.

[67] H. Farid and S. Lyu. Higher-order wavelet statistics and their application to digital forensics. In *Conference on Computer Vision and Pattern Recognition Workshop*, volume 8, page 94, 2003.

[68] X. Feng, I. J. Cox, and G. Doërr. An energy-based method for the forensic detection of re-sampled images. In *IEEE International Conference on Multimedia and Expo*, pages 1–6, 2011.

[69] X. Feng, I.J. Cox, and G. Doërr. Normalized energy density-based forensic detection of resampled images. *IEEE Transactions on Multimedia*, 14(3):536–545, 2012.

[70] X. Feng and G. Doërr. JPEG recompression detection. *SPIE Conference on Media Forensics and Security*, 2010.

[71] P. Ferrara, T. Bianchi, A. De Rosa, and A. Piva. Image forgery localization via fine-grained analysis of CFA artifacts. *IEEE Transactions on Information Forensics and Security*, 7(5):1566–1577, 2012.

[72] C. S. Fillion and G. Sharma. Detecting content adaptive scaling of images for forensic applications. In *SPIE Conference on Media Forensics and Security*, 2010.

[73] R. W. Fleming, A. Torralba, and E. H. Adelson. Specular reflections and the perception of shape. *Journal of Vision*, 4(9), 2004.

[74] M. Fontani, A. Costanzo, M. Barni, T. Bianchi, A. De Rosa, and A. Piva. Two decision fusion frameworks for image forensics. In *Annual GTTI Meeting*, 2011.

[75] J. Fridrich. Digital image forensics using sensor noise. *IEEE Signal Processing Magazine*, 26(2):26–37, 2009.

[76] J. Fridrich. Sensor defects in digital image forensics. In Husrev Taha Sencar and Nasir Memon, editors, *Digital Image Forensics*, pages 179–218. Springer New York, 2013.

[77] J. Fridrich, D. Soukal, and J. Lukáš. Detection of copy move forgery in digital images. In *Digital Forensic Research Workshop*, 2003.

[78] D. Fu, Y. Q. Shi, and W. Su. Detection of image splicing based on hilbert-huang transform and moments of characteristic functions with wavelet decomposition. In YunQing Shi and Byeungwoo Jeon, editors, *Digital Watermarking*, volume 4283 of *Lecture Notes in Computer Science*, pages 177–187. Springer Berlin Heidelberg, 2006.

[79] D. Fu, Y. Q. Shi, and W. Su. A generalized Benford's law for JPEG coefficients and its applications in image forensics. In Edward J. Delp and Ping Wah Wong, editors, *SPIE Conference on Security, Steganography, and Watermarking of Multimedia Contents*, volume 6505, 2007.

[80] A. C. Gallagher. Detection of linear and cubic interpolation in JPEG compressed images. In *Second Canadian Conference on Computer and Robot Vision*, pages 65–72, 2005.

[81] A. C. Gallagher and T. Chen. Image authentication by detecting traces of demosaicing. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pages 1–8, 2008.

[82] T. Gloe, M. Kirchner, A. Winkler, and R. Böhme. Can we trust digital image forensics? In *Proceedings of the 15th international conference on Multimedia*, pages 78–86, New York, NY, USA, 2007. ACM.

[83] M. Goljan and J. Fridrich. Camera identification from scaled and cropped images. In Edward J. Delp and Ping Wah Wong, editors, *SPIE Conference on Security, Forensics, Steganography, and Watermarking of Multimedia Contents*, volume 6819, 2008.

[84] M. Goljan and J. Fridrich. Sensor-fingerprint based identification of images corrected for lens distortion. *Proceedings of SPIE Media Watermarking, Security, and Forensics*, 2012.

[85] M. Goljan, J. Fridrich, and M. Chen. Sensor noise camera identification: Countering counter-forensics. In *SPIE Conference on Media Forensics and Security*, 2010.

[86] M. Goljan, J. Fridrich, and M. Chen. Defending against fingerprint-copy attack in sensor-based camera identification. *IEEE Transactions on Information Forensics and Security*, 6(1):227–236, 2011.

[87] E. S. Gopi. Digital image forgery detection using artificial neural network and independent component analysis. *Applied Mathematics and Computation*, 194(2):540–543, 2007.

[88] E. S. Gopi, L. Nataraj, T. Gokul, S. KumaraGanesh, and P. R. Shah. Digital image forgery detection using artificial neural network and auto regressive coefficients. In *Canadian Conference on Electrical and Computer Engineering*, pages 194–197, 2006.

[89] H. Gou, A. Swaminathan, and M. Wu. Noise features for image tampering detection and steganalysis. In *IEEE International Conference on Image Processing*, volume 6, pages VI–97–VI–100, 2007.

[90] J. Grim, P. Somol, and P. Pudil. Digital image forgery detection by local statistical models. In *International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pages 579–582, 2010.

[91] K. Hara, K. Nishino, and K. Ikeuchi. Light source position and reflectance estimation from a single view without the distant illumination assumption. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(4):493–505, 2005.

[92] J. He, Z. Lin, L. Wang, and X. Tang. Detecting doctored JPEG images via DCT coefficient analysis. In Aleš Leonardis, Horst Bischof, and Axel Pinz, editors, *European Conference on Computer Vision*, volume 3953 of *Lecture Notes in Computer Science*, pages 423–435. Springer Berlin Heidelberg, 2006.

[93] D.R. Hougen and N. Ahuja. Estimation of the light source distribution and its use in integrated shape recovery from stereo and shading. In *Proceedings of the 4th International Conference on Computer Vision*, pages 148–155, 1993.

[94] Y. Hsu and S. Chang. Detecting image splicing using geometry invariants and camera characteristics consistency. In *IEEE International Conference on Multimedia and Expo*, 2006.

[95] Y. Hsu and S. Chang. Image splicing detection using camera response function consistency and automatic segmentation. In *IEEE International Conference on Multimedia and Expo*, 2007.

[96] Y. Hsu and S. Chang. Camera response functions for image forensics: An automatic algorithm for splicing detection. *IEEE Transactions on Information Forensics and Security*, 5(4):816–825, 2010.

[97] F. Huang, J. Huang, and Y. Q. Shi. Detecting double JPEG compression with the same quantization matrix. *IEEE Transactions on Information Forensics and Security*, 5(4):848–856, 2010.

[98] H. Huang, W. Guo, and Y. Zhang. Detection of copy-move forgery in digital images using SIFT algorithm. In *Pacific-Asia Workshop on Computational Intelligence and Industrial Application*, pages 272–276, 2008.

[99] Y. Huang and N. Fan. Learning from interpolated images using neural networks for digital forensics. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 177–182, 2010.

[100] Y. Huang and Y. Long. Demosaicking recognition with applications in digital photo authentication based on a quadratic pixel correlation model. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008.

[101] Y. Huang, W. Lu, W. Sun, and D. Long. Improved DCT-based detection of copy-move forgery in images. *Forensic Science International*, 206(1):178–184, 2011.

[102] P. Irawan. *Appearance of woven cloth*. PhD thesis, Cornell University, Ithaca, NY, USA, 2008.

[103] J. Jacobson and S. Werner. Why cast shadows are expendable: Insensitivity of human observers and the inherent ambiguity of cast shadows in pictorial art. *Perception*, 33(11):1369–1383, 2004.

[104] W. Jakob. Mitsuba renderer, 2010. http://www.mitsuba-renderer.org.

[105] M. K. Johnson and H. Farid. Exposing digital forgeries by detecting inconsistencies in lighting. In *ACM Multimedia and Security Workshop*, New York, NY, 2005.

[106] M. K. Johnson and H. Farid. Exposing digital forgeries through chromatic aberration. In *ACM Multimedia and Security Workshop*, pages 48–55, 2006.

[107] M. K. Johnson and H. Farid. Exposing digital forgeries in complex lighting environments. *IEEE Transactions on Information Forensics and Security*, 3(2):450–461, 2007.

[108] M. K. Johnson and H. Farid. Exposing digital forgeries through specular highlights on the eye. In *International Workshop on Information Hiding*, pages 311–325, 2007.

[109] M. K. Johnson and H. Farid. Detecting photographic composites of people. In Yun Q. Shi, Hyoung-Joong Kim, and Stefan Katzenbeisser, editors, *Digital Watermarking*, Lecture Notes in Computer Science, pages 19–33. Springer, 2008.

[110] M. K. Johnson, D. G. Stork, S. Biswas, and Y. Furuichi. Inferring illumination direction estimated from disparate sources in paintings: an investigation into Jan Vermeer's Girl with a Pearl Earring. *Proceedings of SPIE Computer Image Analysis in the Study of Art*, 2008.

[111] P. Kakar and S. Natarajan. Exposing postprocessed copy–paste forgeries through transform-invariant features. *IEEE Transactions on Information Forensics and Security*, 7(3):1018–1028, 2012.

[112] P. Kakar, S. Natarajan, and W. Ser. Image authentication by motion blur consistency verification. In *IEEE Region 10 Conference*, pages 188–193, 2010.

[113] X. Kang and S. Wei. Identifying tampered regions using singular value decomposition in digital image forensics. In *International Conference on Computer Science and Software Engineering*, pages 926–930, 2008.

[114] K. Karsch, V. Hedau, D. Forsyth, and D. Hoiem. Rendering synthetic objects into legacy photographs. In *Proceedings of the 2011 SIGGRAPH Asia Conference*, SA '11, pages 157:1–157:12, New York, NY, USA, 2011. ACM.

[115] Y. Ke, R. Sukthankar, and L. Huston. An efficient parts-based near-duplicate and sub-image retrieval system. In *Proceedings of the 12th annual ACM international conference on Multimedia*, MULTIMEDIA '04, pages 869–876, New York, NY, USA, 2004. ACM.

[116] E. Kee and H. Farid. Digital image authentication from thumbnails. In *SPIE Conference on Media Forensics and Security*, 2010.

[117] E. Kee and H. Farid. Exposing digital forgeries from 3-D lighting environments. In *IEEE International Workshop on Information Forensics and Security*, Seattle, WA, 2010.

[118] E. Kee, M. K. Johnson, and H. Farid. Digital image authentication from JPEG headers. *IEEE Transactions on Information Forensics and Security*, 6(3):1066–1075, 2011.

[119] B. Khang, J. J. Koenderink, and A. Kappers. Perception of illumination direction in images of 3-D convex objects: Influence of surface materials and light fields. *Perception*, 35(5):625–645, 2006.

[120] N. Khanna, G.T.-C. Chiu, J. P. Allebach, and E. J. Delp. Forensic techniques for classifying scanner, computer generated and digital camera images. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 1653–1656, 2008.

[121] M. Kharrazi, H. T. Sencar, and N. Memon. Blind source camera identification. In *IEEE International Conference on Image Processing*, pages 709–712, 2004.

[122] C. Kim, A. P. Petrov, H. Choh, Y. Seo, and I. Kweon. Illuminant direction and shape of a bump. *Journal of the Optical Society of America*, 15(9):2341–2350, 1998.

[123] M. Kirchner. Fast and reliable resampling detection by spectral analysis of fixed linear predictor residue. In *Proceedings of the 10th ACM workshop on Multimedia and Security*, pages 11–20, New York, NY, USA, 2008. ACM.

[124] M. Kirchner. Efficient estimation of cfa pattern configuration in digital camera images. In *SPIE Conference on Media Forensics and Security*, 2010.

[125] M. Kirchner. Linear row and column predictors for the analysis of resized images. In *ACM Workshop on Multimedia and Security*, pages 13–18, 2010.

[126] M. Kirchner and R. Böhme. Tamper hiding: Defeating image forensics. In Teddy Furon, Francois Cayre, Gwenael Doerr, and Patrick Bas, editors, *International Workshop on Information Hiding*, pages 326–341, 2007.

[127] M. Kirchner and R. Böhme. Hiding traces of resampling in digital images. *IEEE Transactions on Information Forensics and Security*, 3(4):582–592, 2008.

[128] M. Kirchner and R. Böhme. Synthesis of color filter array pattern in digital images. In Edward J. Delp, Jana Dittmann, Nasir Memon, and Ping Wah Wong, editors, *SPIE Conference on Media Forensics and Security*, 2009.

[129] M. Kirchner and J. Fridrich. On detection of median filtering in digital images. In *SPIE Conference on Media Forensics and Security*, 2010.

[130] M. Kirchner and T. Gloe. On resampling detection in re-compressed images. In *IEEE International Workshop on Information Forensics and Security*, pages 21–25, 2009.

[131] J. J. Koenderink, S. Pont, A. Van Doorn, A. Kappers, and J. Todd. The visual light field. *Perception*, 36(11):1595–1610, 2007.

[132] J. J. Koenderink, A. Van Doorn, and A. Kappers. Surface perception in pictures. *Attention, Perception, & Psychophysics*, 52:487–496, 1992. 10.3758/BF03206710.

[133] J. J. Koenderink, A. van Doorn, and S. Pont. Light direction from shad(ow)ed random Gaussian surfaces. *Perception*, 33(12):1405–1420, 2004.

[134] J. Kornblum. Using JPEG quantization tables to identify imagery processed by software. *Digital Investigation*, 5, Supplement(0):S21–S25, 2008.

[135] J. Lalonde and A. Efros. Using color compatibility for assessing image realism. In *IEEE International Conference on Computer Vision*, 2007.

[136] A. Langille and M. Gong. An efficient match-based duplication detection algorithm. In *Canadian Conference on Computer and Robot Vision*, page 64, 2006.

[137] T. Van Lanh, K. Chong, S. Emmanuel, and M. S. Kankanhalli. A survey on digital camera image forensic methods. In *IEEE International Conference on Multimedia and Expo*, pages 16–19, 2007.

[138] B. Li, Y. Q. Shi, and J. Huang. Detecting doubly compressed JPEG images by using mode based first digit features. In *IEEE Workshop on Multimedia Signal Processing*, pages 730–735, 2008.

[139] C. Li and Y. Li. Color-decoupled photo response non-uniformity for digital image forensics. *IEEE Transactions on Circuits and Systems for Video Technology*, 22(2):260–271, 2012.

[140] G. Li, Q. Wu, D. Tu, and S. Sun. A sorted neighborhood approach for detecting duplicated regions in image forgeries based on DWT and SVD. In *IEEE International Conference on Multimedia and Expo*, pages 1750–1753, 2007.

[141] L. Li, J. Xue, X. Wang, and L. Tian. A robust approach to detect tampering by exploring correlation patterns. In Pedro Real, Daniel Diaz-Pernil, Helena Molina-Abril, Ainhoa Berciano, and Walter Kropatsch, editors, *Computer Analysis of Images and Patterns*, volume 6855 of *Lecture Notes in Computer Science*, pages 515–522. Springer Berlin Heidelberg, 2011.

[142] W. Li, Y. Yuan, and N. Yu. Detecting copy-paste forgery of JPEG image via block artifact grid extraction. In *International Workshop on Local and Non-Local Approximation in Image Processing*, 2008.

[143] W. Li, T. Zhang, X. Ping, and E. Zheng. Identifying photorealistic computer graphics using second-order difference statistics. In *International Conference on Fuzzy Systems and Knowledge Discovery*, pages 2316–2319, 2010.

[144] Y. Li, S. Lin, H. Lu, and H. Shum. Multiple-cue illumination estimation in textured scenes. In *Proceedings of the 9th IEEE International Conference on Computer Vision*, pages 1366–1373 vol.2, 2003.

[145] Z. Li and J. Zheng. Blind detection of digital forgery image based on the local entropy of the gradient. In *International Workshop on Digital Watermarking*, pages 161–169, 2008.

[146] C. Lien, C. Shih, and C. Chou. Fast forgery detection with the intrinsic resampling properties. In *International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pages 232–235, 2010.

[147] H. Lin, C. Wang, and Y. Kao. Fast copy-move forgery detection. *WSEAS Transactions on Signal Processing*, 5(5):188–197, 2009.

[148] W. S. Lin, S. Tjoa, H. V. Zhao, and K. J. R. Liu. Digital image source coder forensics via intrinsic fingerprints. *IEEE Transactions on Information Forensics and Security*, 4(3):460–475, 2009.

[149] Z. Lin, J. He, X. Tang, and C. Tang. Fast, automatic and fine-grained tampered JPEG image detection via dct coefficient analysis. *Pattern Recognition*, 42(11):2492–2501, 2009.

[150] Z. Lint, R. Wang, X. Tang, and H. Shum. Detecting doctored images using camera response normality and consistency. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 1087–1092 vol. 1, 2005.

[151] Q. Liu, X. Cao, C. Deng, and X. Guo. Identifying image composites through shadow matte consistency. *IEEE Transactions on Information Forensics and Security*, 6(3):1111–1122, 2011.

[152] Q. Liu and A. H. Sung. A new approach for JPEG resize and image splicing detection. In *ACM Multimedia and Security Workshop*, pages 43–48, 2009.

[153] W. Lu, W. Sun, F. Chung, and H. Lu. Revealing digital fakery using multiresolution decomposition and higher order statistics. *Engineering Applications of Artificial Intelligence*, 24(4):666–672, 2011.

[154] J. Lukáš and J. Fridrich. Estimation of primary quantization matrix in double compressed JPEG images. In *Digital Forensic Research Workshop*, 2003.

[155] J. Lukáš, J. Fridrich, and M. Goljan. Digital camera identification from sensor noise. *IEEE Transactions on Information Forensics and Security*, 1(2):205–214, 2006.

[156] W. Luo, J. Huang, and G. Qiu. Robust detection of region-duplication forgery in digital images. In *International Conference on Pattern Recognition*, pages 746–749, 2006.

[157] W. Luo, Z. Qu, J. Huang, and G. Qiu. A novel method for detecting cropped and recompressed image block. In *IEEE Conference on Acoustics, Speech, and Signal Processing*, pages 217–220, 2007.

[158] S. Lyu. *Natural Image Statistics for Digital Image Forensics*. PhD thesis, Department of Computer Science, Dartmouth College, Hanover, NH, 2005.

[159] S. Lyu. Estimating vignetting function from a single image for image authentication. In *ACM Workshop on Multimedia and Security*, pages 3–12, 2010.

[160] S. Lyu. *Natural Image Statistics in Digital Image Forensics*, pages 239–256. Springer-Verlag, 2013.

[161] S. Lyu and H. Farid. How realistic is photorealistic? *IEEE Transactions on Signal Processing*, 53(2):845–850, 2005.

[162] S. Devi Mahalakshmi, K. Vijayalakshmi, and S. Priyadharsini. Digital image forgery detection and estimation by exploring basic image manipulations. *Digital Investigation*, 8(3-4):215–225, 2012.

[163] B. Mahdian and S. Saic. Detection of copy-move forgery using a method based on blur moment invariants. *Forensic Science International*, 171(2-3):180–189, 2007.

[164] B. Mahdian and S. Saic. Blind authentication using periodic properties of interpolation. *IEEE Transactions on Information Forensics and Security*, 3(3):529–538, 2008.

[165] B. Mahdian and S. Saic. Blind methods for detecting image fakery. In *IEEE International Carnahan Conference on Security Technology*, pages 280–286, 2008.

[166] B. Mahdian and S. Saic. Detection of resampling supplemented with noise inconsistencies analysis for image forensics. In *International Conference on Computational Sciences and Its Applications*, pages 546–556, 2008.

[167] B. Mahdian and S. Saic. A cyclostationarity analysis applied to image forensics. In *IEEE Workshop on Applications of Computer Vision*, pages 389–399, 2009.

[168] B. Mahdian and S. Saic. Detecting double compressed JPEG images. In *International Conference on Imaging for Crime Detection and Prevention*, 2009.

[169] B. Mahdian and S. Saic. Detection and description of geometrically transformed digital images. In Edward J. Delp, Jana Dittmann, Nasir Memon, and Ping Wah Wong, editors, *SPIE Conference on Media Forensics and Security*, 2009.

[170] B. Mahdian and S. Saic. A bibliography on blind methods for identifying image forgery. *Signal Processing: Image Communication*, 25(6):389–399, 2010.

[171] B. Mahdian and S. Stanislav. Using noise inconsistencies for blind image forensics. *Image and Vision Computing*, 27(10):1497–1503, 2009. Special Section: Computer Vision Methods for Ambient Intelligence.

[172] P. Mamassian. Impossible shadows and the shadow correspondence problem. *Perception*, 33(11):1279–1290, 2004.

[173] P. Mamassian and D. Kersten. Illumination, shading and the perception of local orientation. *Vision Research*, 36(15):2351–2367, 1996.

[174] S. R. Marschner and D. P. Greenberg. Inverse lighting for photography. In *Color Imaging Conference*, pages 262–265, 1997.

[175] G. Muhammad, M.. Hussain, K. Khawaji, and G. Bebis. Blind copy move image forgery detection using dyadic undecimated wavelet transform. In *International Conference on Digital Signal Processing*, pages 1–6, 2011.

[176] A. N. Myna, M. G. Venkateshmurthy, and C. G. Patil. Detection of region duplication forgery in digital images using wavelets and log-polar mapping. In *International Conference on Computational Intelligence and Multimedia Applications*, volume 3, pages 371–377, 2007.

[177] L. Nataraj, A. Sarkar, and B. S. Manjunath. Adding Gaussian noise to "denoise" JPEG for detecting image resizing. In *IEEE International Conference on Image Processing*, pages 1493–1496, 2009.

[178] T. Ng and S. Chang. Classifying photographic and photorealistic computer graphic images using natural image statistics. Technical report, ADVENT Technical Report, #220-2006-6, Columbia University, 2004.

[179] T. Ng, S. Chang, Y. Hsu, L. Xie, and M. Tsui. Physics-motivated features for distinguishing photographic images and computer graphics. In *Proceedings of the 13th annual ACM international conference on Multimedia*, pages 239–248, New York, NY, USA, 2005. ACM.

[180] T. Ng, S. Chang, C. Lin, and Q. Sun. *Multimedia Security Technologies for Digital Rights*, chapter Passive-Blind Image Forensics, pages 383–412. Elsevier, 2006.

[181] T. Ng, S. Chang, and Q. Sun. Blind detection of photomontage using higher order statistics. In *International Symposium on Circuits and Systems*, pages 688–691, 2004.

[182] T. Ng, S. Chang, and M. Tsui. Using geometry invariants for camera response function estimation. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2007.

[183] T. Ng and M. Tsui. Camera response function signature for digital forensics — part I: Theory and data selection. In *IEEE Workshop on Information Forensics and Security*, pages 156–160, 2009.

[184] H. C. Nguyen and S. Katzenbeisser. Security of copy-move forgery detection techniques. In *International Conference on Acoustics, Speech and Signal Processing*, pages 1864–1867, 2011.

[185] P. Nillius and J.O. Eklundh. Automatic estimation of the projected light source direction. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages I–1076–I–1083 vol.1, 2001.

[186] J. O'Brien and H. Farid. Exposing photo manipulation with inconsistent reflections. *ACM Transactions on Graphics*, 31(1):4:1–11, 2012. Presented at SIGGRAPH 2012.

[187] T. Okabe, I. Sato, and Y. Sato. Spherical harmonics vs. Haar wavelets: Basis for recovering illumination from cast shadows. volume 1, pages 50–57, Los Alamitos, CA, 2004.

[188] M. Oren and S. Nayar. Generalization of Lambert's Reflectance Model. In *Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pages 239–246, 1994.

[189] J. O'Shea, M. Agrawala, and M. Banks. The influence of shape cues on the perception of lighting directions. *Journal of Vision*, 12(10):21:1–21, 2010.

[190] Y. Ostrovsky, P. Cavanagh, and P. Sinha. Perceiving illumination inconsistencies in scenes. *Perception*, 34(11):1301–1314, 2005.

[191] F. Pan, J. Chen, and J. Huang. Discriminating between photorealistic computer graphics and natural images using fractal geometry. *Science in China Series F: Information Sciences*, 52:329–337, 2009.

[192] F. Pan and J. Huang. Discriminating computer graphics images and natural images using hidden markov tree model. In *International Workshop on Digital Watermarking*, pages 23–28, 2010.

[193] X. Pan and S. Lyu. Region duplication detection using image feature matching. *IEEE Transactions on Information Forensics and Security*, 5(4):857–867, 2010.

[194] X. Pan, X. Zhang, and S. Lyu. Exposing image forgery with blind noise estimation. In *Proceedings of the thirteenth ACM multimedia workshop on Multimedia and security*, MM&Sec'11, pages 15–20, New York, NY, USA, 2011. ACM.

[195] X. Pan, X. Zhang, and S. Lyu. Exposing image splicing with inconsistent local noise variances. In *IEEE International Conference on Computational Photography*, pages 1–10, 2012.

[196] F. Peng, Y. Nie, and M. Long. A complete passive blind image copy-move forensics scheme based on compound statistics features. *Forensic Science International*, 212:e21–e25, 2011.

[197] T. Pevny and J. Fridrich. Detection of double-compression in JPEG images for applications in steganography. *IEEE Transactions on Information Forensics and Security*, 3(2):247–258, 2008.

[198] S. Pfennig and M. Kirchner. Spectral methods to determine the exact scaling factor of resampled digital images. In *International Symposium on Communications Control and Signal Processing*, pages 1–6, 2012.

[199] A. Piva. An overview on image forensics. *ISRN Signal Processing*, 2013:Article ID 496701, 22 pages, 2013.

[200] M. Poilpre, P. Perrot, and H. Talbot. Image tampering detection using bayer interpolation and JPEG compression. In *International Conference on Forensic Applications and Techniques in Telecommunications, Information, and Multimedia*, number 17 in e-Forensics '08, pages 17:1–17:5, ICST, Brussels, Belgium, Belgium, 2008. ICST.

[201] S. Pont and J. Koenderink. Matching illumination of solid objects. *Attention, Perception, & Psychophysics*, 69:459–468, 2007.

[202] A. C. Popescu. *Statistical Tools for Digital Image Forensics*. PhD thesis, Department of Computer Science, Dartmouth College, Hanover, NH, 2005.

[203] A. C. Popescu and H. Farid. Exposing digital forgeries by detecting duplicated image regions. Technical Report TR2004-515, Department of Computer Science, Dartmouth College, 2004.

[204] A. C. Popescu and H. Farid. Exposing digital forgeries by detecting traces of resampling. *IEEE Transactions on Signal Processing*, 53(2):758–767, 2005.

[205] A. C. Popescu and H. Farid. Exposing digital forgeries in color filter array interpolated images. *IEEE Transactions on Signal Processing*, 53(10):3948–3959, 2005.

[206] A. C. Popescu and H. Farid. Statistical tools for digital forensics. In Jessica Fridrich, editor, *Information Hiding*, volume 3200 of *Lecture Notes in Computer Science*, pages 395–407. Springer Berlin / Heidelberg, 2005.

[207] S. Prasad and K. R. Ramakrishnan. On resampling detection and its application to detect image tampering. In *IEEE International Conference on Multimedia and Expo*, pages 1325–1328, 2006.

[208] A. Preetham, P. Shirley, and B. Smits. A practical analytic model for daylight. In *Proceedings of the 26th annual conference on computer graphics and interactive techniques*, pages 91–100. ACM Press/Addison-Wesley Publishing Co., 1999.

[209] Z. Qu, W. Luo, and J. Huang. A convolutive mixing model for shifted double JPEG compression with application to passive image authentication. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 1661–1664, 2008.

[210] R. Ramamoorthi and P. Hanrahan. On the relationship between radiance and irradiance: determining the illumination from images of a convex Lambertian object. *Journal of the Optical Society of America*, 18(10):2448–2459, 2001.

[211] C. Riess and E. Angelopoulou. Scene illumination as an indicator of image manipulation. In Rainer Böhme, Philip W.L. Fong, and Reihaneh Safavi-Naini, editors, *Information Hiding*, volume 6387 of *Lecture Notes in Computer Science*, pages 66–80. Springer Berlin Heidelberg, 2010.

[212] A. Rocha and S. Goldenstein. Is it fake or real? In *Brazilian Symposium of Computer Graphics and Image Processing*, 2006.

[213] A. Rocha, W. Scheirer, T. E. Boult, and S. Goldenstein. Vision of the unseen: Current trends and challenges in digital image and video forensics. *ACM Computing Surveys (CSUR)*, 43(4):26:1–26:42, 2011.

[214] K. Rosenfeld and H. T. Sencar. A study of the robustness of PRNU-based camera identification. In Edward J. Delp, Jana Dittmann, Nasir Memon, and Ping Wah Wong, editors, *SPIE Conference on Media Forensics and Security*, 2009.

[215] S. Ryu, M. Lee, and H. Lee. Detection of copy-rotate-move forgery using Zernike moments. In Rainer Böhme, Philip W.L. Fong, and Reihaneh Safavi-Naini, editors, *Information Hiding*, volume 6387 of *Lecture Notes in Computer Science*, pages 51–65. Springer Berlin Heidelberg, 2010.

[216] P. Saboia, T. Carvalho, and A. Rocha. Eye specular highlights telltales for digital forensics: A machine learning approach. In *International Conference on Image Processing*, pages 1937–1940, 2011.

[217] G. Sankar, H. V. Zhao, and Y. Yang. Feature based classification of computer graphics and real images. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 1513–1516, 2009.

[218] A. Sarkar, L. Nataraj, and B. S. Manjunath. Detection of seam carving and localization of seam insertions in digital images. In *ACM Multimedia and Security Workshop*, pages 107–116, 2009.

[219] I. Sato, Y. Sato, and K. Ikeuchi. Illumination from shadows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(3):290–300, 2003.

[220] H. T. Sencar and N. Memon. Overview of state-of-the-art in digital image forensics. *Algorithms, Architectures, and Information Systems Security 3*, pages 325–348, 2008.

[221] A. Shesh, A. Criminisi, C. Rother, and G. Smyth. 3D-aware image editing for out of bounds photography. In *Proceedings of Graphics Interface*, pages 47–54, 2009.

[222] Y. Q. Shi, C. Chen, and W. Chen. A natural image model approach to splicing detection. In *ACM Multimedia and Security Workshop*, pages 51–62, 2007.

[223] Y. Q. Shi, C. Chen, G. Xuan, and W. Su. Steganalysis versus splicing detection. In YunQ. Shi, Hyoung-Joong Kim, and Stefan Katzenbeisser, editors, *Digital Watermarking*, volume 5041 of *Lecture Notes in Computer Science*, pages 158–172. Springer Berlin Heidelberg, 2008.

[224] M. Sorell. Conditions for effective detection and identification of primary quantization of re-quantized JPEG images. In *International Conference on Forensic Applications and Techniques in Telecommunications, Information, and Multimedia*, number 18 in e-Forensics '08, pages 18:1–18:6, ICST, Brussels, Belgium, Belgium, 2008. ICST.

[225] M. Stamm and K. J. R. Liu. Blind forensics of contrast enhancement in digital images. In *IEEE International Conference on Image Processing*, pages 3112–3115, 2008.

[226] M. Stamm and K. J. R. Liu. Forensic detection of image tampering using intrinsic statistical fingerprints in histograms. In *Proceeding of Asia-Pacific Signal and Information Processing Association, 2009 Annual Summit and Conference*, pages 563–572, 2009.

[227] M. Stamm and K. J. R. Liu. Forensic detection of image manipulation using statistical intrinsic fingerprints. *IEEE Transactions on Information Forensics and Security*, 5(3):492–506, 2010.

[228] M. Stamm and K. J. R. Liu. Forensic estimation and reconstruction of a contrast enhancement mapping. In *International Conference on Acoustics, Speech, and Signal Processing*, pages 1698–1701, 2010.

[229] M. Stamm and K. J. R. Liu. Anti-forensics of digital image compression. *IEEE Transactions on Information Forensics and Security*, 6(3):1050–1065, 2011.

[230] M. Stamm, S. Tjoa, W. S. Lin, and K. J. R. Liu. Anti-forensics of JPEG compression. In *International Conference on Acoustics, Speech, and Signal Processing*, pages 1694–1697, 2010.

[231] M. Stamm, S. Tjoa, W. S. Lin, and K. J. R. Liu. Undetectable image tampering through JPEG compression anti-forensics. In *International Conference on Image Processing*, pages 2109–2112, 2010.

[232] D. G. Stork. Did Georges de la tour use optical projections while painting Christ in the Carpenter's Studio? *Proceedings of SPIE Image and Video Communications and Processing*, pages 214–219, 2005.

[233] D. G. Stork. Mathematical foundations for quantifying shape, shading, and cast shadows in realist master drawings and paintings. *Proceedings of SPIE Mathematics of Data/Image Pattern Recognition, Compression, and Encryption with Applications IX*, pages 63150K–63150K–6, 2006.

[234] D. G. Stork and M. K. Johnson. Computer vision, image analysis, and master art: Part 2. *IEEE MultiMedia*, 13(4):12–17, 2006.

[235] D. G. Stork and M. K. Johnson. Estimating the location of illuminants in realist master paintings computer image analysis addresses a debate in art history of the Baroque. In *International Conference on Pattern Recognition*, volume 1, pages 255–258, 2006.

[236] D. G. Stork and M. K. Johnson. Lighting analysis of diffusely illuminated tableaus in realist paintings: an application to detecting "compositing" in the portraits of Garth Herrick. *Proceedings of SPIE Media Forensics and Security*, 2009.

[237] Y. Sutcu, B. Coskun, H. T. Sencar, and N. Memon. Tamper detection based on regularity of wavelet transform coefficients. In *IEEE International Conference on Image Processing*, pages 24–27, 2007.

[238] P. Sutthiwan, J. Ye, and Y. Q. Shi. An enhanced statistical approach to identifying photorealistic images. In AnthonyT.S. Ho, YunQ. Shi, H.J. Kim, and Mauro Barni, editors, *Digital Watermarking*, volume 5703 of *Lecture Notes in Computer Science*, pages 323–335. Springer Berlin Heidelberg, 2009.

[239] A. Swaminathan, M. Wu, and K. J. R. Liu. Component forensics of digital cameras: A non-intrusive approach. In *Information Sciences and Systems*, pages 1194–1199, 2006.

[240] A. Swaminathan, M. Wu, and K. J. R. Liu. Image tampering identification using blind deconvolution. In *IEEE International Conference on Image Processing*, pages 2309–2312, 2006.

[241] A. Swaminathan, M. Wu, and K. J. R. Liu. A component estimation framework for information forensics. In *Workshop on Multimedia Signal Processing*, pages 397–400, 2007.

[242] A. Swaminathan, M. Wu, and K. J. R. Liu. Nonintrusive component forensics of visual sensors using output images. *IEEE Transactions on Information Forensics and Security*, 2(1):91–106, 2007.

[243] A. Swaminathan, M. Wu, and K. J. R. Liu. Digital image forensics via intrinsic fingerprints. *IEEE Transactions on Information Forensics and Security*, 3(1):101–117, 2008.

[244] G. Valenzise, V. Nobile, M. Tagliasacchi, and S. Tubaro. Countering JPEG anti-forensics. In *International Conference on Image Processing*, pages 1949–1952, 2011.

[245] G. Valenzise, M. Tagliasacchi, and S. Tubaro. The cost of JPEG compression anti-forensics. In *International Conference on Acoustics, Speech and Signal Processing*, pages 1884–1887, 2011.

[246] J. Wang, G. Liu, H. Li, Y. Dai, and Z. Wang. Detection of image region duplication forgery using model with circle block. In *International Conference on Multimedia Information Networking and Security*, pages 25–29, 2009.

[247] J. Wang and H. Zhang. Exposing digital forgeries by detecting traces of image splicing. In *International Conference on Signal Processing*, volume 2, pages 16–20, 2006.

[248] W. Wang, J. Dong, and T. Tan. Effective image splicing detection based on image chroma. In *IEEE International Conference on Image Processing*, pages 1257–1260, 2009.

[249] W. Wang, J. Dong, and T. Tan. Exploring DCT coefficient quantization effect for image tampering localization. In *Workshop on Information Forensics and Security*, pages 1–6, 2011.

[250] W. Wang and H. Farid. Detecting re-projected video. In *International Workshop on Information Hiding*, pages 72–86, 2008.

[251] X. Wang, B. Xuan, and S. Peng. Digital image forgery detection based on the consistency of defocus blur. In *International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pages 192–195, 2008.

[252] Y. Wang and P. Moulin. On discrimination between photorealistic and photographic images. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 2, page II, 2006.

[253] Y. Wang and D. Samaras. Estimation of multiple directional light sources for synthesis of augmented reality images. *Graphical Models*, 65(4):185–205, 2003.

[254] W. Wei, S. Wang, and Z. Tang. Estimation of rescaling factor and detection of image splicing. In *International Conference on Communication Technology*, pages 676–679, 2008.

[255] W. Wei, S. Wang, X. Zhang, and Z. Tang. Estimation of image rotation angle using interpolation-related spectral signatures with application to blind detection of image forgery. *IEEE Transactions on Information Forensics and Security*, 5(3):507–517, 2010.

[256] R. Wu, X. Li, and B. Yang. Identifying computer generated graphics via histogram features. In *IEEE International Conference on Image Processing*, pages 1933–1936, 2011.

[257] T. Yamasaki, T. Matsunami, and K. Aizawa. Detecting resized JPEG images by analyzing high frequency elements in DCT coefficients. In *International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pages 567–570, 2010.

[258] S. Ye, Q. Sun, and E. Chang. Detecting digital image forgeries by measuring inconsistencies of blocking artifact. In *IEEE International Conference on Multimedia and Expo*, pages 12–15, 2007.

[259] I. Yerushalmy and H. Hel-Or. Digital image forgery detection based on lens and sensor aberration. *International Journal of Computer Vision*, 92(1):71–91, 2011.

[260] H. Yuan. Blind forensics of median filtering in digital images. *IEEE Transactions on Information Forensics and Security*, 6(4):1335–1345, 2011.

[261] F. Zach, C. Riess, and E. Angelopoulou. Automated image forgery detection through classification of JPEG ghosts. In *Pattern Recognition, Joint 34th DAGM and 36th OAGM Symposium*, pages 185–194, 2012.

[262] C. Zhang and H. Zhang. Exposing digital image forgeries by using canonical correlation analysis. In *International Conference on Pattern Recognition*, pages 838–841, 2010.

[263] J. Zhang, Z. Feng, and Y. Su. A new approach for detecting copy-move forgery in digital images. In *IEEE Singapore International Conference on Communication Systems*, pages 362–366, 2008.

[264] W. Zhang, X. Cao, Z. Feng, J. Zhang, and P. Wang. Detecting photographic composites using two-view geometrical constraints. In *IEEE International Conference on Multimedia and Expo*, pages 1078–1081, 2009.

[265] W. Zhang, X. Cao, J. Zhang, J. Zhu, and P. Wang. Detecting photographic composites using shadows. In *IEEE International Conference on Multimedia and Expo*, pages 1042–1045, 2009.

[266] Y. Zhang and Y.-H. Yang. Multiple illuminant direction detection with application to image synthesis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(8):915–920, 2001.

[267] E. Zheng, X. Ping, and T. Zhang. Detecting composite JPEG images in transform domain. In *International Conference on Multimedia Information Networking and Security*, pages 340–344, 2010.

[268] Q. Zheng and R. Chellappa. Estimation of illuminant direction, albedo, and shape from shading. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 540–545, 1991.

[269] W. Zhou and C. Kambhamettu. A unified framework for scene illuminant estimation. *Image and Vision Computing*, 26(3):415–429, 2008.