Color Features and Local Structure in Images

Joost van de Weijer

This book is typeset by the author using LaTeX2$_\varepsilon$.

Color Features and Local Structure
in Images

Promotiecommissie:

Promotor: Prof. dr. ir. A. W. M. Smeulders
Co-promotor: dr. T. Gevers

Overige leden: Prof. dr. ir. E. Granum
                Prof. dr. ir. F. C. A. Groen
                Prof. dr. J. J. Koenderink
                Prof. dr. ir. L. J. van Vliet
                Dr. J. M. Geusebroek

Faculteit: Natuurwetenschappen, Wiskunde & Informatica

Netherlands Organisation for Scientific Research

Advanced School for Computing and Imaging
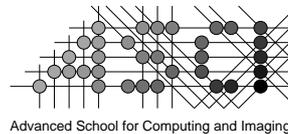
Intelligent Sensory Information Systems
University of Amsterdam
The Netherlands

Contents

Chapter 1

Introduction

When asked about the importance of color, Picasso in one of his blue years exclaimed: "Colors are only symbols. Reality is to be found in luminance alone." His message seems to be taken to heart by the computer vision community. In general the first thing to do, when trying to interpret the content of images, when looking for objects, persons, textures, or at a smaller scale for edges, ridges, and corners, is to discard color. In fact, color is seen as superfluous in a world which can be very well understood by considering luminance alone. This is reflected in the fact that the majority of the current existing computer vision applications is solely based on luminance.

When asking a person who became colorblind later in life, about the importance of color. He will answer that he sometimes wrongly identifies objects where there are only shadows present. When driving the car he suddenly brakes to stop for a shadow blocking the road [61]. Next to that, he sometimes encounters problems distinguishing between objects, e.g. mistaking ketchup for jam, and mustard for mayonnaise [61]. These confusions caused by color blindness surely point out the importance of color in interpreting the visible world.

Two major advantages of using color vision are revealed from the previous example. First, color provides extra information which allows the distinction between various physical causes for color variations in the world, such as changes due to shadows, light source reflections, and object reflectance variations. This helps to quickly identify the black object on the road as a shadow. Next to this, color is an important discriminative property of objects, allowing us to distinguish between mustard and mayonnaise. This thesis explores these aspects of color, proposing theory and techniques to improve the usefulness of color for computer vision.

1.1 Color in Computer Vision

1.1.1 From Luminance to Color

From a mathematical viewpoint the extension from luminance to color signals is an extension from scalar-signals to vector-signals. This change is accompanied by
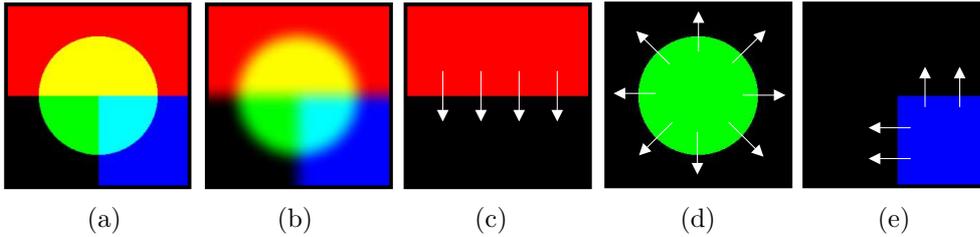
Figure 1.1: *(a) Example image and (b) linear smoothed version of example image. (c) Red channel, (d) green channel and (e) blue channel of example image (see also color plate C.1).*

several mathematical obstacles. Straightforward application of existing luminance-based operators on the separate color channels, and subsequent combination of the results, fails due to undesired artifacts [1].

For example, smoothing a color image with a Gaussian filter blurs the edges, which is also common for luminance based smoothing. In color images linear smoothing introduces new chromaticities. An example is given in Fig. 1.1 where, after linear smoothing, the color purple appears between the blue and red region. These new chromaticities are visually unacceptable and new techniques are required for the task of color image enhancement.

To prevent the introduction of new chromaticities, non-linear operations are required. In contrast to luminance values there is no natural ordering for vector values, meaning that there is no generally accepted method to say that one vector is larger than another. Therefore, new algorithms are required for the computation of known non-linear operators such as the median, local and global mode [1], [59], [73], [84]. A framework in which these non-linear operators are elegantly brought together is the Imprecision Space of Griffin [22], also known as locally orderless images by Koenderink [41], [84]. Apart from the spatial scale, defining the size of the spatial extent of a measurement, the tonal scale is introduced describing its extent along the intensity axis. As a consequence, points are no longer described by a single value, but by a local histogram instead. Extension of this framework to color, although straightforward, is practically unusable due to the computational complexity caused by the high-dimensionality of such color histograms. Since the operations based on the local histograms remain desired for color images, efficient algorithms are needed which prevent the actual computation of the local color histograms.

A second mathematical hurdle in the extension from luminance to color-based operations is how to combine the differential structure of color images. Combining the derivatives with a simple addition of the separate channels results in cancellation in the case of opposing vectors [11]. This is illustrated in Fig. 1.1c,d,e. For the blue-red and cyan-yellow edge in Fig. 1.1 the vectors in the red and blue channel point in opposite directions and a summation will result in a zero edge response, while an edge is obviously present. Also for more complex local features, such as corners and T-junctions, the combination of the channels poses problems. Applying a

(a)   (b)   (c)
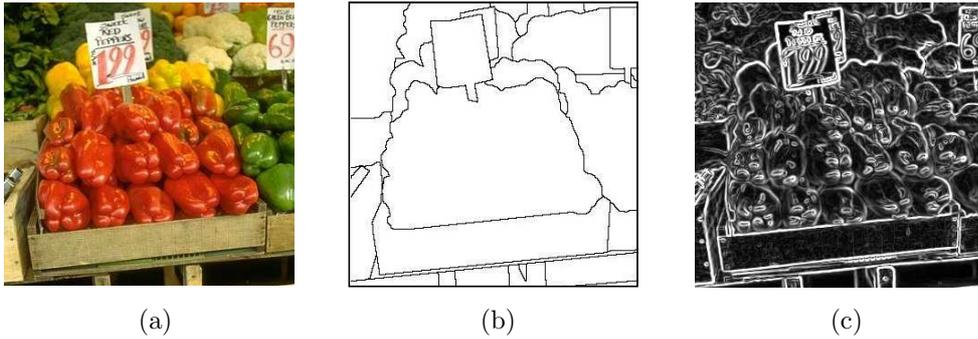
Figure 1.2: *(a) Example image, (b) human scene segmentation and (c) standard computer edge detection (see also color plate C.2).*

corner detector to the separate channels results in a single detected corner in the blue channel. However, there is no evidence for the cross-points with the circle in any of the separate channels. Hence, a combination of corner information from the separate channels might fail. New methods are required to combine the differential structure of color images in a principled way.

1.1.2   Photometric Information

There are several causes of color value composition in images, including shadows, shading, specularity and material edges. In Fig. 1.2, an example of a real-world scene is given, together with a human segmentation (groundtruth) [50]. Furthermore, in Fig. 1.2c the result of a standard edge detection algorithm is given. The algorithm returns more edges than the human segmentation. The problem is how to measure the importance of edges. An important indicator may be derived from the physical cause of an edge. Is the edge caused by a shadow, shading, highlight, or a object reflectance change in the scene? The human segmentation discards all scene incidental edges, such as the shading of the peppers and the specularities. Hence, for scene interpretation it is important to distinguish between the various causes of features in images.

The dichromatic reflection model, introduced to computer vision by Shafer [66], provides a physical model which identifies how photometric changes influence RGB-values. Based on this model, others provided methods for segmentation, classification, and recognition which are independent of scene incidental events. These methods focussed on zeroth order photometric invariance [18], [20], [38], [52], [71]. The effect of the dichromatic model on higher order, differential-based algorithms remained unexplored for long.

Differential photometric invariance is investigated by Geusebroek et al. [16]. The drawbacks of photometric invariance theory, loss of discriminative power and deterioration of noise characteristics [69], are inherited by the differential photometric operations. To improve performance, the impact of the instability of photometric invariants can be diminished through a noise propagation analysis of the invariants [19].

(a)   (b)

Figure 1.3: *(a) Example image, and (b) results of a standard salient point detector (see also color plate C.3).*

However, a drawback is that proper noise estimation is required which is not always available. Hence, methods are required to compute robust photometric invariants without a-priori knowledge of the noise characteristics.

1.1.3   Color Distinctiveness

Describing objects in the world as a set of salient points is currently used with success in object recognition, matching and retrieval [47], [63], [65], [85]. The distinctiveness of the selected salient points is of critical importance for the applicability of the method. It defines the conciseness of the representation and the discriminative power of the local features.

   For example, in Fig. 1.3, a picture of two brightly colored parrots on a dull background is depicted. In Fig. 1.3b the most prominent corners computed by the Harris salient point detector are depicted [27]. Only four out of the twenty points correspond to the salient parrots. And, none of the points focus on the bright red-yellow transition which immediately attracts the eye. The weakness of the salient point detector is mainly in its disregard of color information.

   Although the majority of image data is in color format nowadays only little work has been done in incorporating color into salient point detection and evaluation [29], [33]. One of the reasons luminance-based methods remain much used is because the lack of significant improvement with respect to luminance based methods. This can be explained by the important observation that the majority of differential variation in color images is along the luminance axis. A drawback of the success of the luminance representation, is that when looking for rare events, the axis of major variation is of much less importance. For the computation of the distinctive points in the image, the focus should be on rare events. For these events the axes of relatively little variation become indispensable, and hence for salient point detection color information is crucial.

1.2 Objectives and Approach

In this thesis, we aim to improve the three aspects of color vision discussed above. From the above discussion we arrive at the following three objectives:

1. **From Luminance to Color:** Extend luminance-based algorithms to color in a mathematically sound way. One consequence is that color image enhancement methods do not introduce new chromaticities. A second implication is that for differential-based algorithms the derivatives of the separate channels should be combined without loss of derivative information.

In chapter 2, an image enhancement method for color images is proposed which is based on the minimization of a robust error norm [9], [32]. Interpreting color image enhancement as a robust estimation problem reduces the introduction of unwanted new chromaticities. In the case of a zeroth order local model, the method is proven to be equal to finding the local mode in a histogram. However, it has the advantage that the histogram is never computed. Higher order local models allow for more complex local structures, and therefore yield better image enhancement results.

The problem of opposing vectors, which occurs for all color image edges, only occurs for a particular class of luminance images. Namely, for oriented patterns, which are patterns with one dominant orientation, such as fingerprint data and seismic images [37]. These patterns are characterized by their high frequency nature. The local differential structure consists of quickly succeeding valleys and ridges, with local gradients pointing in opposing directions. Existing operations fail on these images since they are designed for neighborhoods which can be locally modelled as a step-edge. To cope with the opposing vector problem new operations are needed. The solution is found in tensor mathematics, in which opposing vectors reinforce each other [6], [8], [24]. For local curvature estimation the existing method [86] also fails for oriented patterns. In chapter 4, tensor mathematics is used to derive a local curvature estimator for oriented patterns.

Due to the symmetry between the opposing vector problem for color images and oriented patterns, the operations which were proposed for oriented pattern images are straightforwardly extendable to color images. In chapter 5, an overview of tensor-based features [6], [27], [44] is given and extensions of the features to color images are proposed.

We focus on low-level operations when incorporating color into existing luminance-based algorithms. To handle the mathematical obstacles two methods are proposed. Firstly, for color image enhancement a method is proposed which prevents the introduction of new chromaticities. Secondly, a mathematical model is proposed which combines the differential structure of the color channels.

2. **Photometric Information:** Compute photometric invariant differential information in a robust way. Here we focus on the class of applications for which no a-priori knowledge of the noise characteristics of the acquisition system is available.

In chapter 3, a new set of derivatives is proposed which we refer to as quasi-invariants. These quasi-invariants share with full photometric invariants [16], [18] the property that they are insensitive to certain photometric edges, but do not have the inherent instabilities of full photometric invariants.

In chapter 5, a framework for color image features is proposed which couples color tensor-based features with photometric quasi-invariants and full photometric invariants. The applicability of the quasi-invariants is restricted to feature detection, which is the localization of features in the image. For photometric invariant feature extraction, where local descriptors are extracted from the image, full invariance is still required. To improve the robustness of the full invariants, uncertainty measures of full invariants are derived [19]. The tensor framework elegantly allows incorporation of uncertainty measures. A variety of local image features is derived from this robustified invariant color tensor.

3. **Color Distinctiveness:** Improve the distinctiveness of salient point detection algorithms by explicitly incorporating color statistics into the detector design.

In chapter 6, color distinctiveness is explicitly incorporated into the design of differential-based saliency detection [6], [27], [29]. An algorithm is proposed, which is called color saliency boosting. It starts from an analysis of the statistics of color image derivatives. Based on this study, the salient point detector is adapted in such a way that derivatives with equal saliency have equal impact on the saliency function. The adaptation is general. It is easily extendable to existing feature detectors.

Chapter 2

Least Squares and Robust Estimation of Local Image Structure *

2.1 Introduction

Linear scale-space theory of vision not only refers to the introduction of an explicit scale-parameter, it also refers to the use of differential operators to study the local structure of images. The classical way to observe the local differential image structure is to consider all Gaussian derivatives at scale s up to order N. Basically what we do is construct the Taylor series expansion of the *smoothed* image (i.e. the image observed at scale s). The Taylor polynomial thus is an approximation of the smoothed image and not of the original image.

Instead of constructing a polynomial local model of the smoothed image we can equally well construct a polynomial approximation of the unsmoothed image. Our starting point is the *image facet* model as introduced by Haralick et al. [25]. His facet model takes a polynomial function and fits it to the data observed in a small neighborhood in the image using a linear least squares estimation procedure. The image derivatives then can be calculated as the derivatives of the fitted analytical function.

Farnebäck [12] generalizes the Haralick facet model to incorporate spatial weights in order to express the relative importance of the image samples in estimating the parameters of the polynomial function. In the classic Haralick facet model all points in the local neighborhood are considered equally important.

For spatial weighting the choice of the Gaussian kernel leads to a specially efficient implementation. Due to the fact that the derivatives of the Gaussian function are given by a polynomial (determined by the order of differentiation) times the Gaussian

---

function itself, the coefficients in the polynomial function turn out to be a linear combination of the Gaussian derivatives.

The least squares estimation procedure considers all points in a local neighborhood, even in the situation where the local neighborhood is on the boundary of two regions in an image. The regions on either side of the boundary may well be approximated with a low-order polynomial model. The regions can be so different that their union cannot be accurately described using the same low order polynomial model. The estimation procedure then compromises between the two regions: the edge will be smoothed.

In 2.2 we generalize the Gaussian facet model to deal with those multi-model situations. Instead of using a linear least squares estimation procedure we will use a robust estimation technique. A robust estimation technique will only consider the data points from one of the regions and will disregard the data from the other region as being statistical outliers. Robust estimation of local image structure is pioneered by Besl [5]. Our work (see also [78]) differs from the work of Besl in that we consider Gaussian aperture instead of 'crisp' neighborhoods in which the polynomial function is fitted. Furthermore we introduce a fixed point iteration procedure to find the robust estimate.

In 2.3 we present a generalization of earlier work [78], [81], [82]. We derive iterative robust estimators of local image structure and we will give some examples ranging from a simple zero order Gaussian facet model to a first order facet model for color images.

In 2.4 we describe a robust estimator for a derived image quantity: the local orientation (see also [82]). To that end we consider the often used orientation estimator based on a eigen analysis of the structure tensor. Robust estimation of the orientation turns out to be quite similar, the structure tensor is replaced with a 'robustified' version in which only the points are considered that closely fit the model (i.e. the points that are not outliers).

2.2   Least Squares Estimation of Local Image Structure

Locally around a point $\mathbf{x}$ the image function f can be approximated with a linear combination of basis functions ϕ_i, $i = 1, \ldots, K$:

$$\hat{f} = a_1\phi_1 + \cdots + a_K\phi_K. \tag{2.1}$$

We can rewrite this as $\hat{f} = \Phi\mathbf{a}$ where $\Phi = (\phi_1\,\phi_2\cdots\phi_K)$ and $\mathbf{a} = (a_1\,a_2\cdots a_K)^{\mathsf{T}}$. The least squares estimator minimizes the difference ϵ of the image f and the approximation $\hat{f}$:

$$\epsilon(\mathbf{x}) = \int_{\mathbb{R}^d} \left(f(\mathbf{x} + \mathbf{y}) - \hat{f}(\mathbf{y}) \right)^2 W(\mathbf{y})d\mathbf{y} \tag{2.2}$$

where W is the aperture function defining the locality of the model fitting. Note that the optimal fitting function $\hat{f}$ differs from position to position in the image plane. We thus have that $\hat{f}(\mathbf{y}) = \Phi(\mathbf{y})\mathbf{a}(\mathbf{x})$, i.e. $\hat{f}(\mathbf{y}) = a_1(\mathbf{x})\phi_1(\mathbf{y}) + \cdots + a_K(\mathbf{x})\phi_K(\mathbf{y})$.

The optimal parameter vector $\mathbf{a}$ is found by projecting the function f onto the subspace spanned by the basis functions in Φ. In this function space the inner product is given by:

$$f^\mathsf{T} g \equiv \langle f, g \rangle_W = \int_{\mathbb{R}^d} f(\mathbf{x})\, g(\mathbf{x})\, W(\mathbf{x})\, d\mathbf{x}. \tag{2.3}$$

The inner product of functions f and g will also be denoted as $f^\mathsf{T} g$.

To derive the optimal parameter vector $\mathbf{a}$ we take the derivative of the error ϵ with respect to the parameter vector $\mathbf{a}$, set it equal to zero and solve for $\mathbf{a}$. Writing ϵ in terms of the inner product results in

$$\epsilon(\mathbf{x}) = (f_{-\mathbf{x}} - \Phi \mathbf{a})^\mathsf{T} (f_{-\mathbf{x}} - \Phi \mathbf{a}) \tag{2.4}$$

where $f_{-\mathbf{x}}(\mathbf{y}) = f(\mathbf{x} + \mathbf{y})$ is the translated image $f_{-\mathbf{x}}(\mathbf{y}) = f(\mathbf{x} + \mathbf{y})$. The integral is now 'hidden' in the inner product of two functions. This can be rewritten as:

$$\epsilon(\mathbf{x}) = f_{-\mathbf{x}}^\mathsf{T} f_{-\mathbf{x}} - 2\mathbf{a}^\mathsf{T} \Phi^\mathsf{T} f_{-\mathbf{x}} + \mathbf{a}^\mathsf{T} \Phi^\mathsf{T} \Phi \mathbf{a}. \tag{2.5}$$

Taking the derivative of ϵ with respect to $\mathbf{a}$ and setting this equal to 0 and solving for $\mathbf{a}$ we obtain:

$$\mathbf{a} = (\Phi^\mathsf{T} \Phi)^{-1} \Phi^\mathsf{T} f_{-\mathbf{x}} = \tilde{\Phi}^\mathsf{T} f_{-\mathbf{x}} \tag{2.6}$$

where $\tilde{\Phi} = \Phi(\Phi^\mathsf{T}\Phi)^{-1}$ is the *dual basis*. The functions in the dual basis, $\tilde{\Phi} = \left(\tilde{\phi}_1 \cdots \tilde{\phi}_K \right)$, are the functions such that the inner product $\tilde{\phi}_i^\mathsf{T} f_{-\mathbf{x}}$ equals the coefficient a_i in the approximation $\hat{f} = a_1 \phi_1 + \cdots + a_K \phi_K$. The dual basis functions, multiplied with the aperture function, thus are the correlation kernels needed to calculate the coefficients in the polynomial image approximation.

The classic Haralick facet model uses a uniform weight function $W(\mathbf{x}) = 1$ for $\|\mathbf{x}\|_\infty \leq s$ and $W(\mathbf{x}) = 0$ elsewhere, i.e. a 'crisp' neighborhood within an axis aligned square of size $2s \times 2s$.

For the second order polynomial basis:

$$\Phi = \left(\begin{array}{cccccc} 1, & x, & y, & \frac{1}{2}x^2, & xy, & \frac{1}{2}y^2 \end{array} \right) \tag{2.7}$$

the dual basis is

$$\tilde{\Phi} = \left(\begin{array}{cccccc} \frac{7}{8\,s^2} - \frac{15\,x^2}{16\,s^4} - \frac{15\,y^2}{16\,s^4}, & \frac{3\,x}{4\,s^4}, & \frac{3\,y}{4\,s^4}, & \frac{-15}{8\,s^4} + \frac{45\,x^2}{8\,s^6}, & \frac{9\,x\,y}{4\,s^6}, & \frac{-15}{8\,s^4} + \frac{45\,y^2}{8\,s^6} \end{array} \right). \tag{2.8}$$

The dual basis functions are depicted in Fig. 2.1. The first dual basis function (multiplied with the aperture function) is the correlation kernel needed to calculate the coefficient of the constant basis function in the approximation of the local image patch. Observe that in the Haralick facet model, the first dual basis function is not everywhere positive. Fig. 2.1 also shows the discrete dual basis functions, these follow from a formulation of the facet model in a discrete image space as can be found in the work of Haralick.

Within a scale-space context the most natural choice is to start with a polynomial basis and a Gaussian aperture function $W = G^s$ where G^s is the Gaussian function
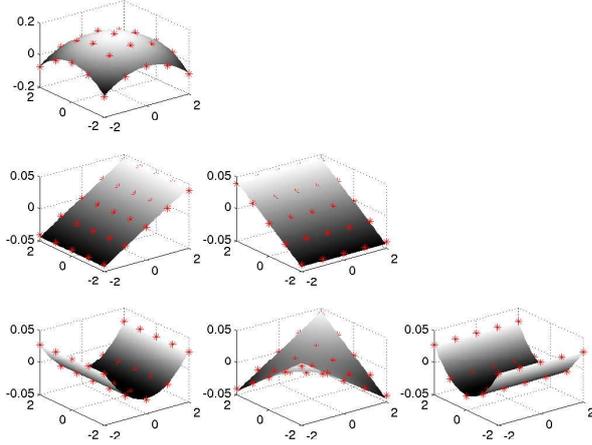
Figure 2.1: **The Haralick Facet Model.** *From left to right, top to bottom the dual basis functions are plotted. The shaded functions are the dual basis functions within a 2nd order facet model, the (red) stars correspond with the discrete dual functions. The neighborhood was taken to be of size 5×5. The scale s for the analytical kernel was set at $s = 2.42$. This value is the value to make the difference between the discrete and analytical facet models minimal. For larger neighborhoods $N \times N$ the correspondence becomes better and the analytical scale approaches $N/2$.*

at scale s. Again starting with the second order polynomial basis the dual basis is a different one due to the difference in the inner product (as a consequence of a different aperture function):

$$\tilde{\Phi} = \left(\quad 2 - \frac{x^2}{2\,s^2} - \frac{y^2}{2\,s^2}, \quad \frac{x}{s^2}, \quad \frac{y}{s^2}, \quad -s^{-2} + \frac{x^2}{s^4}, \quad \frac{x\,y}{s^4}, \quad -s^{-2} + \frac{y^2}{s^4} \quad \right). \qquad (2.9)$$

Again, a dual basis function, multiplied with the—Gaussian—aperture function is the correlation kernel needed to calculate the corresponding coefficient in the polynomial approximation of the local image patch. For the zero order coefficient the correlation kernel is a Gaussian function multiplied with a parabola: $(2 - \frac{x^2}{2\,s^2} - \frac{y^2}{2\,s^2})\,G^s(x,y)$. Again we see that the zero order coefficient in the polynomial image approximation requires a kernel with negative values.

The derivatives of the Gaussian function are equal to a polynomial function (a Hermite polynomial depending on the derivative taken) times the Gaussian function, we may write the correlation kernels associated with the dual basis functions in the Gaussian facet model as a linear combination of Gaussian derivatives. It is not hard to prove that the zero order coefficient in the second order Gaussian facet model is found by convolving the image f with the kernel:

$$G^s - \tfrac{1}{2}s^2 \left(G^s_{xx} + G^s_{yy} \right). \qquad (2.10)$$

Now we easily recognize where the negative values in the kernel come from. The term G^s is the Gaussian scale-space smoothing term. The term $-\frac{1}{2}s^2 \left(G^s_{xx} + G^s_{yy} \right)$ is a
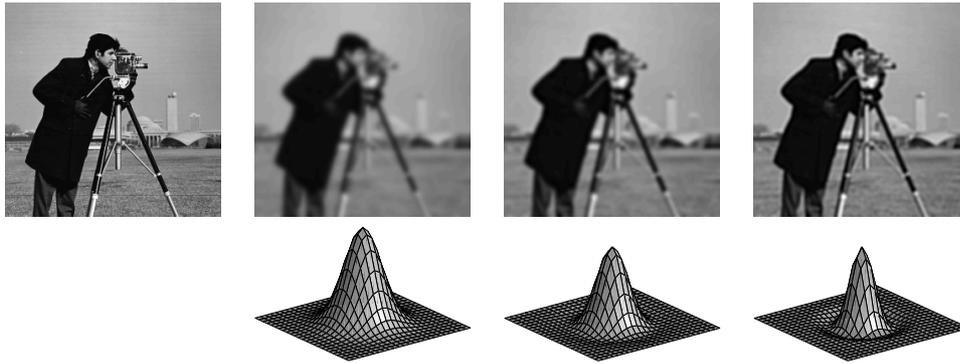
Figure 2.2: **Zero-order coefficient in the Gaussian Facet Model.** *On the first row, from left to right: the original image, and the zero order coefficients in the Gaussian facet model of order 0, 2 and 6. On the second row the convolution kernel is shown that, convoluted with the original image, results in the image above it.*

well-known sharpening term: subtracting the Laplacian from the smoothed image, sharpens the image. The sharpening term is due to the fact that the Gaussian facet model approximates the original image, not the smoothed image.

It turns out that this observation is true for higher order facet models as well. For a 4th order Gaussian facet model, the kernel to calculate the zero order coefficient is:

$$G^s - \tfrac{1}{2}s^2\left(G^s_{xx} + G^s_{yy}\right) + \tfrac{1}{8}s^4\left(G^s_{xxxx} + 2\,G^s_{xxyy} + G^s_{yyyy}\right). \qquad (2.11)$$

In Fig. 2.2 the kernels to calculate the zero order coefficient in the Gaussian facet model of orders 0, 2 and 6 are depicted together with the convoluted images. Apparently the N-jet of an image observed at scale s encodes details of size less then s, i.e. from the N-jet observed at scale s a lot of detail can be reconstructed.

2.3 Robust Estimation of Local Image Structure

Consider again the error of the Gaussian weighted least squares approximation:

$$\epsilon(\mathbf{x}) = \int_{\mathbb{R}^d} \left(f(\mathbf{x} + \mathbf{y}) - \hat{f}(\mathbf{y})\right)^2 G^s(\mathbf{y})d\mathbf{y}. \qquad (2.12)$$

It is well known that this error definition is not well suited for those situations were we have outliers in our measurements. In the image processing context statistical outliers are not so frequently occurring. The effect that makes least squares estimates questionable is that when collecting measurements from a neighborhood in an image these are often not well modelled using a simple (facet) model. For instance we may model local image luminance quite well with a second order polynomial model *but not near edges* where we switch from one model instantiation to another. Such multi-model situations are abundant in computer vision applications and are most often
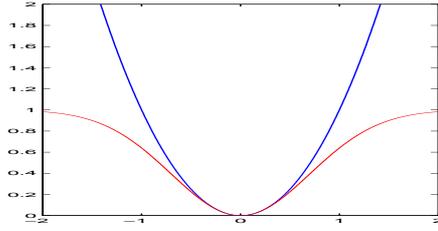
Figure 2.3: **Quadratic versus (robust) Gaussian error norm.** *The Gaussian error norm is of 'scale' $m = 0.7$.*

due to the nature of the imaging process where we see abrupt changes going from one object to another object.

Multi-modality can be incorporated into sophisticated estimation procedures where we not only estimate (multi-)model parameters but also the geometry that separates the different regions (one for each model). One of the oldest examples is perhaps Hueckels edge detector [30] in which a local image patch is described with two regions separated by a straight boundary. The detector estimates this boundary and the parameters of the luminance distributions on each side of the edge.

In this paper we take a less principled approach. Instead of a multi-model approach we stick to a simpler one-model approach where we use a statistical *robust estimator* that allows us to consider part of the measurements from the local neighborhood to belong to the model we are interested in and disregard all other measurements as being 'outliers' and therefore not relevant in estimating the model parameters.

The crux of a robust estimation procedure is to rewrite the above error measure as:

$$\epsilon(\mathbf{x}) = \int_{\mathbb{R}^d} \rho(f(\mathbf{x} + \mathbf{y}) - \hat{f}(\mathbf{y})) \, G^s(\mathbf{y}) d\mathbf{y} \tag{2.13}$$

where ρ is the error norm. The choice $\rho(e) = e^2$ leads to the least squares estimator. Evidently measurements that are outliers to the 'true' model are weighted heavily in the total error measure. Reducing the influence of the large errors leads to *robust error norms*.

Writing $f_{-\mathbf{x}}(\mathbf{y}) = f(\mathbf{x} + \mathbf{y})$ and using the local linear model $\hat{f}(\mathbf{y}) = \Phi(\mathbf{y})\mathbf{a}(\mathbf{x})$ we obtain:

$$\epsilon(\mathbf{x}) = \int_{\mathbb{R}^d} \rho(f_{-\mathbf{x}} - \Phi\mathbf{a}(\mathbf{x})) \, G^s \, d\mathbf{y}. \tag{2.14}$$

We omitted the spatial argument $\mathbf{y}$ for ease of notation. In this paper the 'Gaussian error norm' is chosen:

$$\rho(e) = 1 - \exp\left(-\frac{e^2}{2m^2}\right). \tag{2.15}$$

The scale m in the error norm will be called the *model scale* to contrast it with the spatial scale s that is used in the spatial aperture function G^s. In Fig. 2.3 the error norm is sketched. Compared to the quadratic error norm this norm is 'clamped' at

value 1. For $e \gg m$ the exact value of the error is not important any more. Gross outliers are therefore not given the weight to influence the estimation greatly.

The optimal model parameters are found by calculating the derivative of the error measure and setting this equal to zero:

$$
\begin{aligned}
\frac{\partial \epsilon}{\partial \mathbf{a}} &= \frac{\partial}{\partial \mathbf{a}} \int_{\mathbb{R}^d} \rho(f_{-\mathbf{x}} - \Phi \mathbf{a}(\mathbf{x})) \, G^s \, d\mathbf{y} &\text{(2.16)} \\
&= \frac{\partial}{\partial \mathbf{a}} \int_{\mathbb{R}^d} \left(1 - \exp\left(-\frac{(f_{-\mathbf{x}} - \Phi \mathbf{a}(\mathbf{x}))^2}{2\,m^2} \right) \right) G^S \, d\mathbf{y} &\text{(2.17)} \\
&= -\frac{1}{m} \int_{\mathbb{R}^d} (f_{-\mathbf{x}} - \Phi \mathbf{a}(\mathbf{x})) \, \Phi \exp\left(-\frac{(f_{-\mathbf{x}} - \Phi \mathbf{a}(\mathbf{x}))^2}{2\,m^2} \right) G^s \, d\mathbf{y}. &\text{(2.18)}
\end{aligned}
$$

Setting this derivative equal to zero and rewriting terms we obtain:

$$
\int_{\mathbb{R}^d} f_{-\mathbf{x}} \, \Phi \exp\left(-\frac{(f_{-\mathbf{x}} - \Phi \mathbf{a}(\mathbf{x}))^2}{2\,m^2} \right) G^s \, d\mathbf{y} =
$$
$$
\int_{\mathbb{R}^d} \Phi \mathbf{a}(\mathbf{x}) \, \Phi \exp\left(-\frac{(f_{-\mathbf{x}} - \Phi \mathbf{a}(\mathbf{x}))^2}{2\,m^2} \right) G^s \, d\mathbf{y}. \quad \text{(2.19)}
$$

This can be rewritten as:

$$
\int_{\mathbb{R}^d} f_{-\mathbf{x}} \, \Phi G^m \left(f_{-\mathbf{x}} - \Phi \mathbf{a}(\mathbf{x}) \right) G^s \, d\mathbf{y} = \int_{\mathbb{R}^d} \Phi \mathbf{a}(\mathbf{x}) \, \Phi G^m \left(f_{-\mathbf{x}} - \Phi \mathbf{a}(\mathbf{x}) \right) G^s \, d\mathbf{y} \quad \text{(2.20)}
$$

where G^m is the Gaussian function at scale m. This Gaussian function weighs the model distance, whereas the Gaussian function G^s weighs the spatial distance.

We define the operator Γ:

$$
(\Gamma^m g)(\mathbf{y}) = G^m(f_{-\mathbf{x}}(\mathbf{y}) - \Phi(\mathbf{y})\mathbf{a}(\mathbf{x}))) \, g(\mathbf{y}) \quad \text{(2.21)}
$$

i.e. the point wise multiplication of the function g with the model weight function. Now Γ^m acts as a diagonal (matrix) operator in the function space. Using the vectorial notation of the inner product we can write:

$$
\Phi^{\mathsf{T}} \Gamma^m f_{-\mathbf{x}} = \Phi^{\mathsf{T}} \Gamma^m \Phi \, \mathbf{a}. \quad \text{(2.22)}
$$

This looks like a familiar weighted linear least squares equation that can be solved for the value of $\mathbf{a}$. It is not, because Γ^m is dependent on $\mathbf{a}$. Solving for $\mathbf{a}$ can be done using an *iterated weighted least squares* procedure:

$$
\mathbf{a}^{i+1} = \left(\Phi^{\mathsf{T}} \Gamma(\mathbf{a}^i) \Phi \right)^{-1} \Phi^{\mathsf{T}} \Gamma(\mathbf{a}^i) f_{-\mathbf{x}}. \quad \text{(2.23)}
$$

Some examples of these robust estimators may clarify matters. In the next subsection we consider the most simple of all local structure models: a locally constant model. The resulting image operator turns out to be an iterated version of the bilateral filter introduced by Tomasi and Manduchi [72].
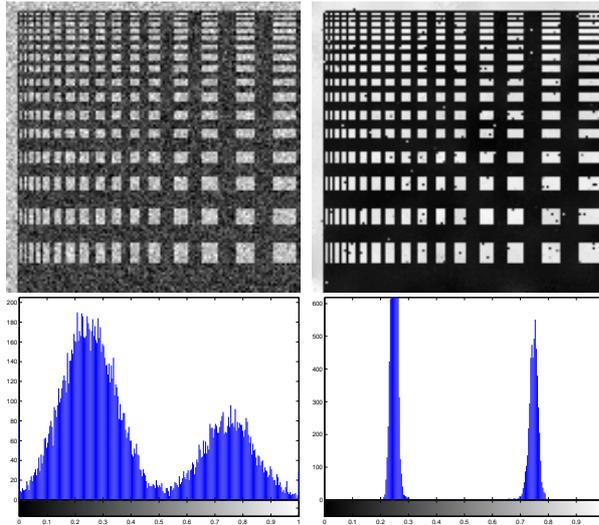
Figure 2.4: **Robust Estimation of Local Image Structure.** *On the first row a test image with noise added on the left and the result of the robust estimator based on a zero-order facet model. On the second row the histograms of the images above are depicted. Observe that the robust estimator is capable of finding the modes of both the distributions.*

2.3.1 Zero-order Image Structure

Consider a locally constant image model with only one basis function:

$$\Phi = (1) \tag{2.24}$$

i.e. the constant function. Eq.(2.23) then reduces to:

$$a_0^{i+1}(\mathbf{x}) = \frac{\int_{\mathbb{R}^d} f(\mathbf{x}+\mathbf{y})\, G^m(f(\mathbf{x}+\mathbf{y}) - a_0^i(\mathbf{x}))\, G^s(\mathbf{y})\, d\mathbf{y}}{\int_{\mathbb{R}^d} G^m(f(\mathbf{x}+\mathbf{y}) - a_0^i(\mathbf{x}))\, G^s(\mathbf{y})\, d\mathbf{y}}. \tag{2.25}$$

This is an iterated version of the bilateral filter as introduced by Tomasi and Manduchi [72]. It is also related to the filters introduced by Smith et al. [70]. The bilateral filter thus implements one iteration of a robust estimator with initial value $a_0^0 = f$.

In previous papers [78], [81] we have analyzed robust estimation of the zero order local image structure. Some observations made are:

- The robust estimator finds the local mode in the local luminance histogram which is smoothed with a Gaussian kernel of scale m. The local mode that is found is the local maximum in the smoothed histogram that is closest to the initial value.

- Bilateral filtering implements one iteration of the robust estimator. From mean shift analysis we know that the first step in a mean shift algorithm is a large one

Figure 2.5: **Robust Estimation and Non-linear diffusion.** *On the left the original image of a flower. In the middle the robust estimation of the zero order local structure and on the right the result of iteratively applying one iteration of the robust estimator, each time using the image data from the previous iteration (this procedure is very much like a non-linear diffusion process).*

in the direction of the optimal value. This explains the impressive results on the bilateral filter in reducing the noise while preserving the structure of images.

- The choice of an initial estimate is very important. We have found good results using the result of a linear least squares estimate as the initial estimate. In certain situations however the amount of smoothing induced by the least squares estimator sets the robust estimator at a wrong starting point leading to a local maximum in the histogram that does not correspond with the structure that we are interested in. This situation is often occurring in case the area of the structure of interest is less then the area of the 'background' (e.g. document images where there is more paper then ink visible). In such cases the image itself can be used as an initial estimate of the zero order local structure.

- The results of robust estimation of local image structure bear great resemblance to the results of non-linear diffusion. The theoretical link between robust estimation and non-linear diffusion techniques has been reported before (see [9]). The main difference with the robust estimator technique described here is that in each iteration of a non-linear diffusion algorithm the image data resulting from the previous iteration is used. In the robust estimator described here we stick to the original image data and only update the parameter to be estimated. Fig. 2.5 shows the differences between these two procedures.

2.3.2   Higher-order Image Structure

For the image in Fig. 2.4 the assumption of local constant image model is a correct assumption, for most natural images such a model is an oversimplification though. Then it is better to use a higher order model for the local image structure. We start with a simple first order model for 1D functions. The local basis is:

$$\Phi = \begin{pmatrix} 1 & x \end{pmatrix}. \tag{2.26}$$
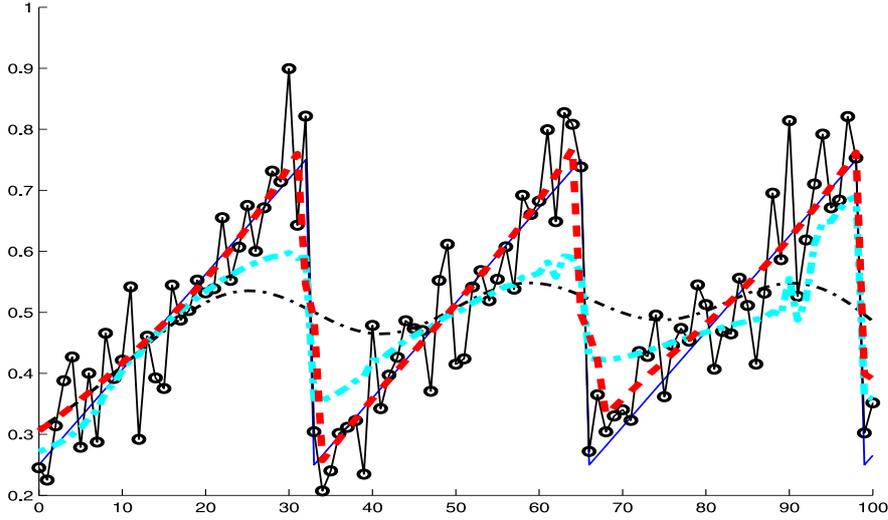
Figure 2.6: **Robust Estimation of Local Structure in 1D functions.** *A 'saw-tooth' function with added noise is shown together with the Gaussian linear least squares estimate, i.e. the Gaussian smoothing (the thin 'sinusoidal' line), the robust estimate based on a zero order facet model (the dashed-dotted line) and the robust estimate based on a first order model (the thick dashed line). The spatial scale is 9 and the tonal (model) scale is 0.1. The number of iterations used is 10.*

This leads to the matrix $\Phi^{\mathsf{T}}\Gamma^m\Phi$:

$$\begin{pmatrix} \int_{\mathbb{R}} G^m(f(x+y) - a_0^i - a_1^i y)G^s(y)dy & \int_{\mathbb{R}} y\, G^m(f(x+y) - a_0^i - a_1^i y)\, G^s(y)\, dy \\ \int_{\mathbb{R}} y\, G^m(f(x+y) - a_0^i - a_1^i y)\, G^s(y)\, dy & \int_{\mathbb{R}} y^2\, G^m(f(x+y) - a_0^i - a_1^i y)\, G^s(y)\, dy \end{pmatrix} \tag{2.27}$$

and vector $\Phi^{\mathsf{T}}\Gamma^m f_{-x}$:

$$\begin{pmatrix} \int_{\mathbb{R}} f(x+y)\, G^m(f(x+y) - a_0^i - a_1^i y)\, G^s(y)\, dy \\ \int_{\mathbb{R}} y\, f(x+y)\, G^m(f(x+y) - a_0^i - a_1^i y)\, G^s(y)\, dy \end{pmatrix}. \tag{2.28}$$

The robust estimator of the local linear model is given by Eq.(2.23). Fig. 2.6 shows a univariate 'saw-tooth' signal corrupted with additive noise. Also shown are the robust estimates based on a zero order facet model and the robust estimate based on a first order facet model. It is obvious that a robust estimator based on a local constant model is not capable of reconstructing the saw tooth signal from the noisy observations. Using a local first order model leads to a far better reconstruction.

The first order robust facet model is easily generalized to 2D functions:

$$\Phi \quad = \quad \begin{pmatrix} \phi_{(00)} & \phi_{(10)} & \phi_{(01)} \end{pmatrix} \tag{2.29}$$

$$= \quad \begin{pmatrix} 1 & x_1 & x_2 \end{pmatrix}. \tag{2.30}$$

Figure 2.7: **Robust Estimation of Local Image Structure.** *On the left the cameraman image with noise added and on the right the robust estimation of the zero order coefficient in a first order facet model.*

This leads to the matrix $\Phi^{\mathsf{T}}\Gamma^m\Phi$:

$$\begin{pmatrix} \int_{\mathbb{R}^2} G^m G^s d\mathbf{y} & \int_{\mathbb{R}^2} y_1 G^m G^s d\mathbf{y} & \int_{\mathbb{R}^2} y_2 G^m G^s d\mathbf{y} \\ \int_{\mathbb{R}^2} y_1 G^m G^s d\mathbf{y} & \int_{\mathbb{R}^2} y_1^2 G^m G^s d\mathbf{y} & \int_{\mathbb{R}^2} y_1 y_2 G^m G^s d\mathbf{y} \\ \int_{\mathbb{R}^2} y_2 G^m G^s d\mathbf{y} & \int_{\mathbb{R}^2} y_1 y_2 G^m G^s d\mathbf{y} & \int_{\mathbb{R}^2} y_2^2 G^m G^s d\mathbf{y} \end{pmatrix} \qquad (2.31)$$

to simplify the notation we have omitted the arguments of the functions in the integrand. For the G^m-function the argument is the model error $f(\mathbf{x}+\mathbf{y}) - a_{00} - a_{10}y_1 - a_{01}y_2$. The vector $\Phi^{\mathsf{T}}\Gamma^m f_{-\mathbf{x}}$ equals

$$\begin{pmatrix} \int_{\mathbb{R}^2} f(\mathbf{x}+\mathbf{y})G^m(f(\mathbf{x}+\mathbf{y}) - a_{00} - a_{10}y_1 - a_{01}y_2)G^s(\mathbf{y})d\mathbf{y} \\ \int_{\mathbb{R}^2} y_1 f(\mathbf{x}+\mathbf{y})G^m(f(\mathbf{x}+\mathbf{y}) - a_{00} - a_{10}y_1 - a_{01}y_2)G^s(\mathbf{y})d\mathbf{y} \\ \int_{\mathbb{R}^2} y_2 f(\mathbf{x}+\mathbf{y})G^m(f(\mathbf{x}+\mathbf{y}) - a_{00} - a_{10}y_1 - a_{01}y_2)G^s(\mathbf{y})d\mathbf{y} \end{pmatrix}. \qquad (2.32)$$

Eq.(2.23) then can be used to calculate the new estimate of the optimal parameter vector $\mathbf{a}^{i+1}$.

In Fig. 2.7 the robust estimation of the zero order coefficient based on a first order facet model is shown. For this image the difference with a zero order facet model estimation can only be observed in regions of slowly varying luminance (like in the background).

2.3.3 Color Image Structure

In this section we generalize the robust facet models for scalar images to models for vectorial images. The analysis is done for color images but is valid for all vectorial images.

A color image $\mathbf{f} = (f^1 \ f^2 \ f^3)$ at any position $\mathbf{x}$ has three color components $f^1(\mathbf{x})$, $f^2(\mathbf{x})$ and $f^3(\mathbf{x})$. The local model for a color image using a basis

$$\Phi = (\phi_1 \ \phi_2 \ \cdots \ \phi_K) \qquad (2.33)$$
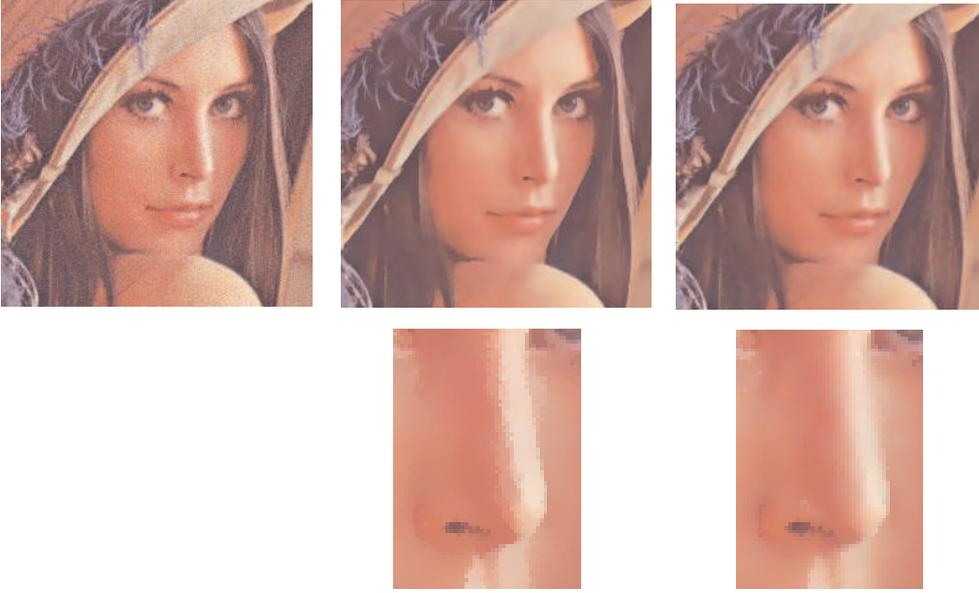
Figure 2.8: **Robust Estimation of Local Structure in Color Images.** *On the first row from left to right: the 'Lena' image with some noise added to it, the zero-order facet model based robust estimator of the values and the robust estimator based on a first order based facet model. On the second row we show a detail from the image above (see also color plate C.4).*

is chosen as:

$$\hat{\mathbf{f}}(\mathbf{x}+\mathbf{y}) = \Phi\mathbf{A} = \Phi \left(\begin{array}{ccc} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{array} \right) \tag{2.34}$$

where $\mathbf{A} = \left(\begin{array}{ccc} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{array} \right)$ is the $K \times 3$ parameter matrix. The column $\mathbf{a}_i$ represents the parameter vector in the approximation $\hat{f}_i = \Phi\mathbf{a}_i$ of the i-th color component. Each of the color components is thus approximated as a linear combination of K basis functions. The model error is now written as:

$$\epsilon(\mathbf{x}) = \int_{\mathbb{R}^d} \rho \left(\sqrt{(f^1_{-\mathbf{x}} - \Phi\mathbf{a}_1)^2 + (f^2_{-\mathbf{x}} - \Phi\mathbf{a}_2)^2 + (f^3_{-\mathbf{x}} - \Phi\mathbf{a}_3)^2} \right) G^s(\mathbf{y})d\mathbf{y}. \tag{2.35}$$

It is not hard to prove that in this case

$$\frac{\partial\epsilon}{\partial\mathbf{A}} = 0 \iff \Phi^{\mathsf{T}}\Gamma^m\mathbf{f} = \Phi^{\mathsf{T}}\Gamma^m\Phi\mathbf{A} \tag{2.36}$$

where Γ^m is the 'diagonal' operator that multiplies a function point wise with the function: $G^m \left((f^1_{-\mathbf{x}} - \Phi\mathbf{a}_1)^2 + (f^2_{-\mathbf{x}} - \Phi\mathbf{a}_2)^2 + (f^3_{-\mathbf{x}} - \Phi\mathbf{a}_3)^2 \right)$. As Γ^m is dependent on the parameter matrix $\mathbf{A}$ we arrive at a iterated weighted least squares estimator:

$$\mathbf{A}^{i+1} = (\Phi^{\mathsf{T}}\Gamma^m(\mathbf{A}^i)\Phi)^{-1}\Phi^{\mathsf{T}}\Gamma^m(\mathbf{A}^i)\,\mathbf{f}. \tag{2.37}$$
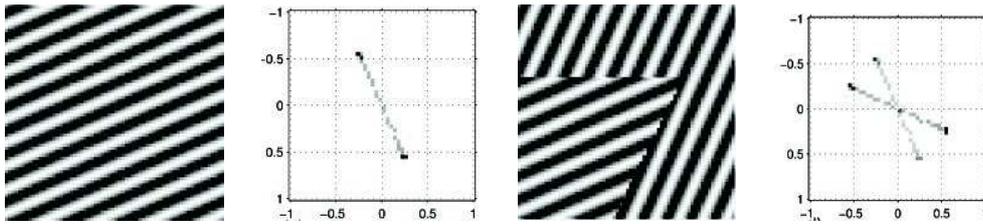
Figure 2.9: **Histograms of gradient vector space.** *In (a) an image (64 × 64) is shown with in (b) the histogram of all gradient vectors (where darker shades indicate that those gradient vectors occur often in the image. In (c) a composition of two differently oriented patterns is shown with corresponding histogram in (d).*

The estimation of the robust facet model for color images is thus almost the same as for scalar images. The three color components are dealt with independently, only the error weights operator Γ^m is dependent on all three color components.

In Fig. 2.8 the robust estimators are shown that are based on a zero order facet model and on a first order facet model. Especially in the nose-region the first order model based robust estimator performs better then the zero order model based robust estimator.

2.4   Robust Estimation of Orientation

In the previous sections we have considered local image models for the image values (grey value and color). In this section we look at robust estimation of the orientation of image structures.

Oriented patterns are found in many imaging applications, e.g. in fingerprint analysis, and in geo-physical analysis of soil layers. The classical technique to estimate the orientation of the texture is to look at the set of luminance gradient vectors in a local neighborhood. In an image patch showing a stripe pattern in only one orientation we can clearly distinguish the orientation as the line cluster in gradient space perpendicular to the stripes (see Fig. 2.9(a-b)). A straightforward eigenvector analysis of the covariance matrix will reveal the orientation of the texture. The covariance matrix of the gradient vectors in an image neighborhood is often used to estimate the local orientation [37], [8], [46], [91].

In case the local neighborhood is taken from the border of two differently oriented patterns (see Fig. 2.9) an eigenvector analysis of the covariance matrix will mix both orientations resulting in a 'smoothing' of the orientation estimation.

In case the regions showing different textures are of sufficient size it is possible to use a post-processing step to sharpen the smoothed orientation measurements. A classical way of doing so is the Kuwahara-Nagao operator [43], [56], [2]. At a certain position in an image this operator searches for a nearby neighborhood where the (orientation) response is more homogeneous then it is at the border. That response is then used at the point of interest. In this way the neighborhoods are not

allowed to cross the borders of the textured regions. In [80] we have shown that the classic Kuwahara-Nagao operator can be interpreted as a 'macroscopic' version of a PDE image evolution that combines linear diffusion (smoothing) with morphological sharpening.

Again consider the texture in Fig. 2.9(a). The histogram of the gradient vectors in this texture patch is shown in Fig. 2.9(b). Let $\mathbf{v}$ be the true orientation vector of the patch, i.e. the unit vector perpendicular to the stripes. In an ideal image patch every gradient vector should be parallel to the orientation $\mathbf{v}$. In practice they will not be parallel. The error of a gradient vector $\mathbf{g}(\mathbf{y})$ observed in a point $\mathbf{y}$ with respect to the orientation $\mathbf{v}(\mathbf{x})$ of an image patch centered at location $\mathbf{x}$ is defined as:

$$e(\mathbf{x}, \mathbf{y}) = \|\mathbf{g}(\mathbf{y}) - (\mathbf{g}(\mathbf{y})^{\mathsf{T}}\mathbf{v}(\mathbf{x}))\mathbf{v}(\mathbf{x})\|. \tag{2.38}$$

The difference $\mathbf{g}(\mathbf{y}) - (\mathbf{g}(\mathbf{y})^{\mathsf{T}}\mathbf{v}(\mathbf{x}))\mathbf{v}(\mathbf{x})$ is the projection of $\mathbf{g}$ on the normal to $\mathbf{v}$. The error $e(\mathbf{x}, \mathbf{y})$ thus measures the perpendicular distance from the gradient vector $\mathbf{g}(\mathbf{y})$ to the orientation vector $\mathbf{v}(\mathbf{x})$. Integrating the squared error over all positions $\mathbf{y}$ using a soft Gaussian aperture for the neighborhood definition we define the total error:

$$\epsilon(\mathbf{x}) = \int_{\Omega} e^2(\mathbf{x}, \mathbf{y}) G^s(\mathbf{x} - \mathbf{y}) d\mathbf{y}. \tag{2.39}$$

The error measure can be rewritten as:

$$\epsilon = \int_{\Omega} \mathbf{g}^{\mathsf{T}}\mathbf{g} G^s d\mathbf{y} - \int_{\Omega} \mathbf{v}^{\mathsf{T}}(\mathbf{g}\mathbf{g}^{\mathsf{T}})\mathbf{v} G^s d\mathbf{y}. \tag{2.40}$$

where we have omitted the arguments of the functions. Minimizing the error thus is equivalent with maximizing:

$$\int_{\Omega} \mathbf{v}^{\mathsf{T}}(\mathbf{g}\mathbf{g}^{\mathsf{T}})\mathbf{v} G^s d\mathbf{y}, \tag{2.41}$$

subject to the constraint that $\mathbf{v}^{\mathsf{T}}\mathbf{v} = 1$. Note that $\mathbf{v}$ is not dependent on $\mathbf{y}$ so that we have to maximize:

$$\mathbf{v}^{\mathsf{T}} \left(\int_{\Omega} (\mathbf{g}\mathbf{g}^{\mathsf{T}}) G^s d\mathbf{y} \right) \mathbf{v} = \mathbf{v}^{\mathsf{T}} \mu^s \mathbf{v} \tag{2.42}$$

where μ^s is the *structure tensor*.

Using the method of Lagrange multipliers to maximize $\mathbf{v}^{\mathsf{T}}\mu^s\mathbf{v}$ subject to the constraint that $\mathbf{v}^{\mathsf{T}}\mathbf{v} = 1$, we need to find an extremum of

$$\lambda(1 - \mathbf{v}^{\mathsf{T}}\mathbf{v}) + \mathbf{v}^{\mathsf{T}}\mu^s\mathbf{v}. \tag{2.43}$$

Differentiating with respect to $\mathbf{v}$ (remember that $d\mathbf{v}^{\mathsf{T}}A\mathbf{v}/d\mathbf{v} = 2A\mathbf{v}$ in case $A = A^{\mathsf{T}}$) and setting the derivative equal to zero results in:

$$\mu^s\mathbf{v} = \lambda\mathbf{v}. \tag{2.44}$$

The 'best' orientation thus is an eigenvector of the structure tensor. Substitution in the quadratic form then shows that we need the eigenvector corresponding to the largest eigenvalue.

The least squares orientation estimation works well in case all gradients in the ensemble of vectors in an image neighborhood all belong to the same oriented pattern. In case the image patch shows two oriented patterns the least squares estimate will mix the two orientations and give a wrong result.

A robust estimator is constructed by introducing the Gaussian error norm once again:

$$\epsilon(\mathbf{x}) = \int_\Omega \rho(e(\mathbf{x}, \mathbf{y})) G^s(\mathbf{x} - \mathbf{y}) d\mathbf{y}. \tag{2.45}$$

In a robust estimator large deviations from the model are not taken into account very heavily. In our application large deviations from the model are probably due to the mixing of two different linear textures (see Fig. 2.9(c-d)).

The error, Eq.(2.45), can now be rewritten as (we will omit the spatial arguments):

$$\epsilon = \int_\Omega \rho\left(\sqrt{\mathbf{g}^\mathsf{T}\mathbf{g} - \mathbf{v}^\mathsf{T}(\mathbf{g}\mathbf{g}^\mathsf{T})\mathbf{v}}\right) G^s d\mathbf{y}. \tag{2.46}$$

Again we use a Lagrange multiplier to minimize subject to the constraint that $\mathbf{v}^\mathsf{T}\mathbf{v} = 1$:

$$\frac{d}{d\mathbf{v}}\left(\lambda(1 - \mathbf{v}^\mathsf{T}\mathbf{v}) + \int_\Omega \rho\left(\sqrt{\mathbf{g}^\mathsf{T}\mathbf{g} - \mathbf{v}^\mathsf{T}(\mathbf{g}\mathbf{g}^\mathsf{T})\mathbf{v}}\right) G^s d\mathbf{y}\right) = 0. \tag{2.47}$$

Using Eq.(2.15) as the error function leads to

$$\eta(\mathbf{v})\mathbf{v} = \lambda\mathbf{v} \tag{2.48}$$

where

$$\eta(\mathbf{v}) = \int_\Omega \mathbf{g}\mathbf{g}^\mathsf{T} G^m\left(\sqrt{\mathbf{g}^\mathsf{T}\mathbf{g} - \mathbf{v}^\mathsf{T}(\mathbf{g}\mathbf{g}^\mathsf{T})\mathbf{v}}\right) G^s d\mathbf{y}. \tag{2.49}$$

The big difference with the least squares estimator is that now the matrix η is dependent on $\mathbf{v}$ (and on $\mathbf{x}$ as well). Note that η can be called a 'robustified' structure tensor in which the contribution of each gradient vector is weighted not only by its distance to the center point of the neighborhood, but also weighted according to its 'distance' to the orientation model. Weickert et al. [92] also introduce a non linear version of the structure tensor that is close in spirit to the robust structure tensor η.

We propose the following *fixed point* iteration scheme to find a solution. Let $\mathbf{v}^i$ be the orientation vector estimate after i iterations. The estimate is then updated as the eigenvector $\mathbf{v}^{i+1}$ of the matrix $\eta(\mathbf{v}^i)$ corresponding to the largest eigenvalue, i.e. we solve:

$$\eta(\mathbf{v}^i)\mathbf{v}^{i+1} = \lambda\mathbf{v}^{i+1}. \tag{2.50}$$

The proposed scheme is a generalization of the well-known fixed point scheme (also called *functional iteration*) to find a solution of the equation $v = F(v)$.

Note that the iterative scheme does not necessarily lead to the *global* minimum of the error. In fact often we are not even interested in that global minimum. Consider for instance the situation of a point in region A (with orientation α_1) that is surrounded by many points in region B (with orientation β). It is not to difficult to imagine a situation where the points of region B outnumber those in region A.
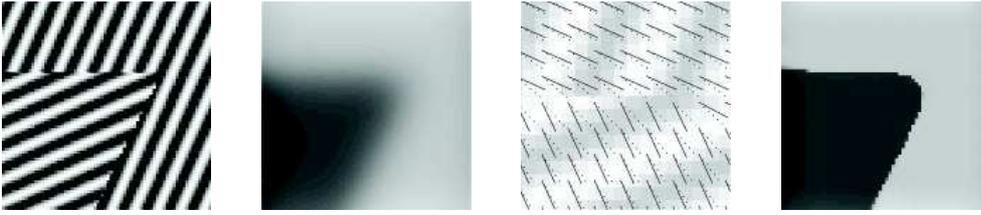
Figure 2.10: **Least Squares versus Robust Orientation Estimation.** *In (a) a generated noise free image is shown. The texture is made out of two regions each differently oriented. In (b) the orientation field $\alpha = \arctan(v_2/v_1)$ is shown that results from the least squares estimate. In (d) the orientation field is shown resulting from the robust estimation. In (c) a detail of the orientation vector fields for both the least squares estimation (dotted lines) and the robust estimation (solid lines) are shown.*
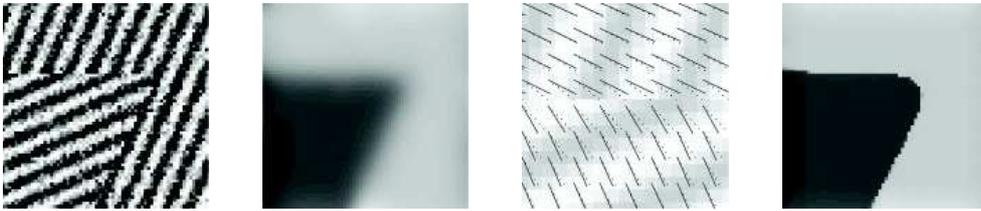
Figure 2.11: **Least Squares versus Robust Orientation Estimation.** *Same experiment as figure 10 but with noise added.*

Nevertheless we would like our algorithm to find the orientation α whereas the global minimum would correspond with orientation β. Because our algorithm starts in the initial orientation estimate and then finds the local minimum nearest to the starting point we hopefully end up in the desired *local* minimum: orientation α.

The choice for an initial estimate of the orientation vector is thus crucial in a robust estimator in case we have an image patch showing multiple striped patterns. In Fig. 2.10 and Fig. 2.11 robust estimation of orientation for a simple test image is given. For the robust estimation we have used the orientation in location **x** that resulted from the least squares estimator as the initial orientation vector in that point. Only 5 iterations are used. For both examples it is evident that the robust estimation performs much better at the border of the textured regions.

Chapter 3

Edge and Corner Detection by Photometric Quasi-Invariants *

3.1 Introduction

Feature detection, such as edge and corner detection, plays an important role in many computer vision applications such as image segmentation, object recognition and image retrieval [26]. A large number of feature detectors is based on the differential structure of images [10], [27], [45]. However, in real-world applications there are various physical phenomena which trigger differential-based features, such as shadows, shading, specularities, and object reflectance changes. It is important to differentiate between the various physical causes of a feature.

An improvement in color understanding was the introduction of the dichromatic reflection model by Shafer [66]. The model separates the reflected light into body reflection (object color) and surface reflection (specularities). This separation results in the classification of physical events, such as shadows and highlights. This is suited for photometric invariant segmentation, object recognition, and retrieval [18], [38], [51]. However, these methods are based on the zeroth order structure of images and mostly involve the analysis of the RGB-values in color histograms. For the photometric invariant theory to be applicable to differential-based operations other methods are needed.

The connection between differential-based features and photometric invariance theory is proposed by Geusebroek et al. [16]. This work provides a set of photometric invariant derivative filters and uses them for invariant edge detection. However, the non-linear transformations used to compute photometric invariants have several draw-

backs such as instabilities and loss of discriminative power. These drawbacks limit the applicability of operations based on derivatives of these invariants. Traditionally, the effect of instabilities is suppressed by ad hoc thresholding of the transformed values [28], [57]. A more elaborate approach is to apply error propagation through the various color spaces to compensate for the undesired effects of instabilities and non-linearities of the different photometric invariant spaces [19]. However, this approach is based on a proper noise estimation system which is not always available in practice.

In this chapter we propose a new class of derivatives which we refer to as photometric quasi-invariants. These derivatives link derivative-based operations to the theory of photometric invariance. Quasi-invariants are derived from the dichromatic reflection model and are proven to differ from full photometric invariants by a scaling factor. These quasi-invariants do not have the inherent instabilities of full photometric invariants, and from theoretical and experimental results it is shown that quasi-invariants have better noise characteristics, discriminative power, and introduce less edge displacement than full photometric invariants. The lack of full photometric invariance limits the applicability of quasi-invariants to methods which are based on a single image, such as edge and corner detection. Quasi-invariants cannot be used for applications in which responses between multiple images are compared, such as invariant object recognition.

3.2   The Dichromatic Reflection Model

In this section the dichromatic reflection model is discussed [66]. The dichromatic model divides the reflection in the body (object color) and surface reflection (specularities or highlights) component for optically inhomogeneous materials. Assuming a known illuminant, $\mathbf{c}^i = (\alpha, \beta, \gamma)^T$, and neutral interface reflection, the RGB vector, $\mathbf{f} = (R, G, B)^T$, can be seen as a weighted summation of two vectors,

$$\mathbf{f} = e(m^b \mathbf{c}^b + m^i \mathbf{c}^i) \tag{3.1}$$

in which $\mathbf{c}^b$ is the color of the body reflectance, $\mathbf{c}^i$ the color of the surface reflectance, m^b and m^i are scalars representing the corresponding magnitudes of body and surface reflection and e is the intensity of the light source. For matte surfaces there is no interface reflection and the model further simplifies to

$$\mathbf{f} = e m^b \mathbf{c}^b \tag{3.2}$$

which is the well-known Lambertian reflection. For more on the validity of the photometric assumptions see [16], [18], [66] and for calibration [19].

From the dichromatic reflection model, photometric invariants can be derived (e.g. normalized RGB, hue). These invariants have the disadvantage that they are unstable; normalized RGB is unstable near zero intensity and hue is undefined on the black-white axis. The instabilities can be avoided by analyzing the RGB values in the RGB-histogram [38] [51]. That proved to be rather difficult and slow since you need a meaningful segmentation to generate a meaningful histogram, and a meaningful histogram to get a good segmentation.
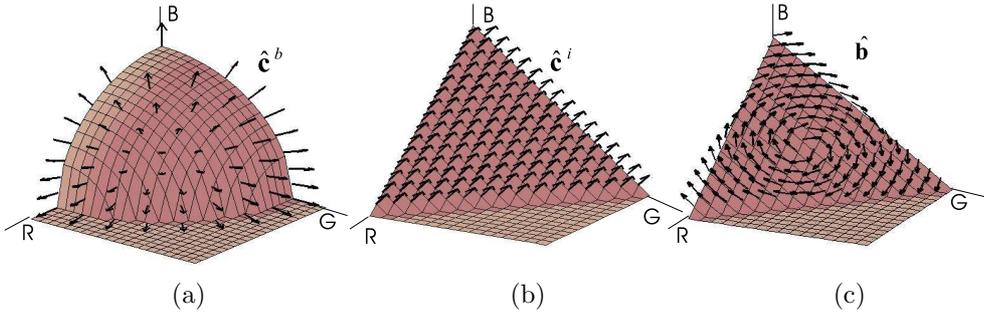
Figure 3.1: *(a) Shadow-shading direction $\hat{\mathbf{c}}^b$, (b) specular direction $\hat{\mathbf{c}}^i$, and (c) hue direction $\hat{\mathbf{b}}$.*

Instead of looking at the zeroth order structure (the RGB-values) we focus in this chapter on the first order structure of the image. A straightforward extension of the photometric invariance theory to first order filters can be obtained by taking the derivative of the invariants. However, these filters would inherit the undesired instabilities of the photometric invariants. Therefore we propose an alternative way to arrive at photometric derivatives by analyzing the spatial derivative of the dichromatic reflection model.

The spatial derivative of the dichromatic reflection model (Eq. 3.1) gives the photometric derivative structure of the image:

$$\mathbf{f}_x = em^b\mathbf{c}_x^b + \left(e_x m^b + em_x^b\right)\mathbf{c}^b + \left(em_x^i + e_x m^i\right)\mathbf{c}^i. \tag{3.3}$$

Here, the subscript indicates spatial differentiation. Since we assume a known illuminant and neutral interface reflection, $\mathbf{c}^i$ is independent of x. The derivative in Eq. 3.3 is a summation of three weighted vectors, successively caused by body reflectance, shading-shadow and specular change. Further, we assume that shadows are not significantly colored.

In fact, the direction of the shadow-shading changes (Fig. 3.1a) follows from Eq. 3.2. In the absence of interface reflection, the direction of $\mathbf{c}^b$ coincides with the direction of $\hat{\mathbf{f}} = \frac{1}{\sqrt{R^2+G^2+B^2}}\left(R, G, B\right)^T$. The hat is used to denote unit vectors. The shadow-shading direction is the multiplication of two scalars denoting two different physical phenomena. First, $e_x m^b$ indicates a change in intensity which corresponds to a shadow edge. And em_x^b is a change in the geometry coefficient which represents a shading edge.

Another direction is the specular direction $\mathbf{c}^i$ in which changes of the specular geometry coefficient m_x^i occur. In Fig. 3.1b, $\mathbf{c}^i$ is depicted for the case of a white light source for which $\hat{\mathbf{c}}^i = \frac{1}{\sqrt{3}}(1, 1, 1)^T$. The specular direction is multiplied by two factors. Firstly, em_x^i is a change of geometric coefficient caused by changes in the angles between viewpoint, object and light source. Secondly, the term $e_x m^i$ representing a shadow edge on top of a specular reflection.

Having the direction of two of the causes of an edge, we are able to construct a third direction which is perpendicular to these two vectors (Fig. 3.1c). This direction,

named hue direction $\hat{\mathbf{b}}$, is computed by the outer product:

$$\hat{\mathbf{b}} = \frac{\hat{\mathbf{f}} \times \hat{\mathbf{c}}^i}{\left| \hat{\mathbf{f}} \times \hat{\mathbf{c}}^i \right|}. \tag{3.4}$$

If $\hat{\mathbf{f}}$ and $\hat{\mathbf{c}}^i$ are parallel, we define $\hat{\mathbf{b}}$ to be the zero vector. Note that the hue direction is *not* equal to the direction in which changes of the body reflectance occur, $\hat{\mathbf{c}}_x^b$. It is perpendicular to the two other causes of an edge. Hence, changes in the hue direction can only be attributed to a body reflectance change.

In conclusion, changes in the reflection manifest themselves as edges in the image. There are three causes for an edge in an image: an hue change, a shadow-shading edge or a specular change. We indicated three directions: the shadow-shading direction, the specular direction and the hue direction. These directions are the same as the directions indicated by Klinker [38] for to use of image segmentation. We use these direction for the construction of photometric invariant spatial derivatives.

3.3 Photometric Variants and Quasi-Invariants

In this section, the goal is to propose a new set of photometric variants and quasi-invariants. To this end, the derivative of an image, $\mathbf{f}_x = (R_x, G_x, B_x)^T$, is projected on three directions found in the previous section. We will call these projections *variants*. E.g. the projection of the derivative on the shadow-shading direction results in the shadow-shading variant. By removing the variance from the derivative of the image, we construct a complementary set of derivatives which we will call *quasi-invariants*.

The projection of the derivative on the shadow-shading direction is called the shadow-shading variant and is defined as

$$\mathbf{S}_x = \left(\mathbf{f}_x \cdot \hat{\mathbf{f}} \right) \hat{\mathbf{f}}. \tag{3.5}$$

The dot indicates the vector inner product. The second $\hat{\mathbf{f}}$ indicates the direction of the variant. The shadow-shading variant is the part of the derivative which could be caused by shadow or shading. Due to correlation of the hue and specular direction with the shadow-shading direction, part of $\mathbf{S}_x$ might be caused by changes in hue or specular reflection.

What remains after subtraction of the variant is called the shadow-shading quasi-invariant, indicated by superscript c,

$$\mathbf{S}_x^c = \mathbf{f}_x - \mathbf{S}_x. \tag{3.6}$$

The quasi-invariant $\mathbf{S}_x^c$ consists of that part of the derivative which is *not* caused by shadow-shading edges (Fig. 3.2b). Hence, only contains specular and hue edges.

The same reasoning can be applied to the specular direction and results in the specular variant and the specular quasi-invariant

$$\begin{aligned} \mathbf{O}_x &= \left(\mathbf{f}_x \cdot \hat{\mathbf{c}}^i \right) \hat{\mathbf{c}}^i \ , \\ \mathbf{O}_x^c &= \mathbf{f}_x - \mathbf{O}_x. \end{aligned} \tag{3.7}$$
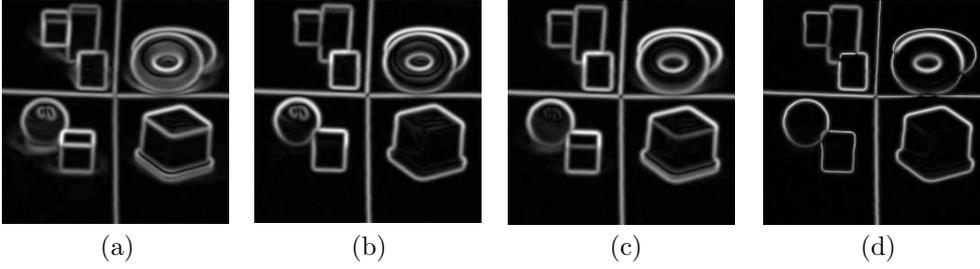
Figure 3.2: *Various derivatives applied to Fig. 3.4a: a) color gradient* ($\mathbf{f}_x$)*, b) shadow-shading quasi-invariant* ($\mathbf{S}_x^c$)*, c) the specular quasi-invariant* ($\mathbf{O}_x^c$) *, and d) the specular-shadow-shading quasi-invariant* ($\mathbf{H}_x^c$)*.*

The specular quasi-invariant is insensitive to highlight edges (Fig. 3.2c).

Finally, we can construct the shadow-shading-specular variant and quasi-invariant by projecting the derivative on the hue direction

$$
\begin{aligned}
\mathbf{H}_x^c &= \left(\mathbf{f}_x \cdot \hat{\mathbf{b}} \right) \hat{\mathbf{b}} \ , \\
\mathbf{H}_x &= \mathbf{f}_x - \mathbf{H}_x^c .
\end{aligned}
\tag{3.8}
$$

$\mathbf{H}_x^c$ does not contain specular or shadow-shading edges (Fig. 3.2d).

3.4 Relations of Quasi-Invariants with Full Invariants

In this section, the resemblances and differences are analyzed between quasi-invariants and full invariants. A geometrical relation in RGB-space between the two is found by investigating underlying color space transformations. Conclusions with respect to stability are made. With stability it is meant that small changes in the RGB-cube do not cause large jumps in the invariant space. Further, we discuss the characteristics of quasi-invariants.

3.4.1 Spherical Color Space

An orthogonal transformation which has the shadow-shading direction as one of its components is the spherical coordinate transformation. Transforming the RGB-color space results in the spherical color space or $r\theta\varphi$-color space. The transformations are,

$$
\begin{aligned}
r &= \sqrt{R^2 + G^2 + B^2} = |\mathbf{f}| \\
\theta &= \arctan(\tfrac{G}{R}) \\
\varphi &= \arcsin\left(\frac{\sqrt{R^2 + G^2}}{\sqrt{R^2 + G^2 + B^2}} \right)
\end{aligned}
\ .
\tag{3.9}
$$

Since r is pointing in the shadow-shading direction, its derivative corresponds to $\mathbf{S}_x$

$$
r_x = \frac{RR_x + GG_x + BB_x}{\sqrt{R^2 + G^2 + B^2}} = \mathbf{f}_x \cdot \hat{\mathbf{f}} = |\mathbf{S}_x| \ .
\tag{3.10}
$$

$$\quad\text{(a)}\qquad\qquad\text{(b)}\qquad\text{(c)}\qquad\qquad\text{(d)}\qquad\qquad\text{(e)}\qquad\text{(f)}$$
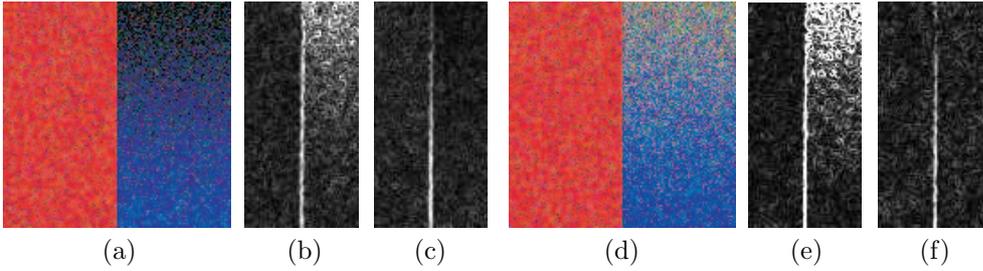
Figure 3.3: *(a) Red-blue edge, with a decreasing intensity of the blue patch going in the upward direction. Response of (b) normalized RGB derivative, and (c) shadow-shading quasi-invariant ($\mathbf{S}_x^c$). (d) Red-blue edge, with decreasing saturation going in the upward direction. Response of (e) hue derivative (h_x), and (f) specular-shadow-shading quasi-invariant ($\mathbf{H}_x^c$) (see also color plate C.5).*

The quasi-invariant $\mathbf{S}_x^c$ is the derivative energy in the plane perpendicular to the shadow-shading direction. The derivative in the $\theta\varphi$-plane is given by

$$\begin{aligned}|\mathbf{S}_x^c| &= \sqrt{(r\varphi_x)^2 + (r\sin\varphi\,\theta_x)^2} \\ &= r\sqrt{(\varphi_x)^2 + (\sin\varphi\,\theta_x)^2}\end{aligned} \quad . \tag{3.11}$$

To conserve the metric of *RGB*-space the angular derivatives are multiplied by their corresponding scale factors which follow from the spherical transformation. For matte surfaces both θ and φ are independent of m^b (substitution of Eq. 3.2 in Eq. 3.9). Hence, the part under the root is a shadow-shading invariant.

By means of the spherical coordinate transformation a relation between the quasi-invariant and the full invariant is found. The difference between the quasi-invariant $|\mathbf{S}_x^c|$ and the full invariant $s_x = \sqrt{(\varphi_x)^2 + (\sin\varphi\,\theta_x)^2}$ is the multiplication with r which is the $L2$ norm for the intensity (see Eq. 3.9). In geometrical terms, the derivative vector which remains after subtraction of the part in the shadow-shading direction is *not* projected on the sphere to produce an invariant. This projection introduces the instability of the full shadow-shading invariants for low intensities,

$$\begin{aligned}&\lim_{r\to 0} s_x \quad \text{does not exist} \\ &\lim_{r\to 0} |\mathbf{S}_x^c| = 0.\end{aligned} \tag{3.12}$$

The first limit follows from the non existence of the limit for both φ_x and θ_x at zero. The second limit can be concluded from $\lim_{r\to 0} r\varphi_x = 0$ and $\lim_{r\to 0} r\theta_x = 0$. Concluding, the multiplication of the full-invariant with $|\mathbf{f}|$ resolves the instability.

An example of the responses for the shadow-shading invariant and quasi-invariant is given in Fig. 3.3. In Fig. 3.3a, a synthetic image of a red-blue edge is depicted. The blue intensity decreases along the y-axis. Gaussian uncorrelated noise is added to the *RGB* channels. In Fig. 3.3b the normalized *RGB* response is depicted and the

instability for low intensities is clearly visible. For the shadow-shading quasi-invariant (Fig. 3.3c), no instability occurs and the response just diminishes for low intensities. Note that the instable region is particularly inconvenient because shadow-shading edges tend to produce low-intensity areas.

3.4.2 Opponent Color Space

The orthonormal transformation which accompanies the specular variant is known as the opponent color space. For a known illuminant $\mathbf{c}^i = (\alpha, \beta, \gamma)^T$ it is given by

$$
\begin{aligned}
o1 &= \frac{\beta R - \alpha G}{\sqrt{\alpha^2 + \beta^2}} \\
o2 &= \frac{\alpha\gamma R + \beta\gamma G - (\alpha^2 + \beta^2)B}{\sqrt{(\alpha^2 + \beta^2 + \gamma^2)(\alpha^2 + \beta^2)}} \\
o3 &= \frac{\alpha R + \beta G + \gamma B}{\sqrt{\alpha^2 + \beta^2 + \gamma^2}}
\end{aligned} \qquad . \tag{3.13}
$$

The relations with the variant and its complement are $|\mathrm{O_x}| = o3_x$ and $|\mathrm{O_x^c}| = \sqrt{o1_x^2 + o2_x^2}$.

3.4.3 The Hue Saturation Intensity Space

As discussed in section 3.3 the shadow-shading-specular quasi-invariant is both perpendicular to the shadow-shading direction and the specular direction. An orthogonal transformation which satisfies this constraint is the hue-saturation-intensity transformation. It is actually a polar transformation on the opponent color axis $o1$ and $o2$.

$$
\begin{aligned}
h &= \arctan\left(\frac{o1}{o2}\right) \\
s &= \sqrt{o1^2 + o2^2} \\
i &= o3
\end{aligned} \qquad . \tag{3.14}
$$

The changes of h occur in the hue direction and hence the derivative in the hue-direction is equal to the shadow-shading-specular quasi-invariant,

$$
|\mathbf{H}_x^c| = s \cdot h_x. \tag{3.15}
$$

The multiplication with the scale factor s follows from the fact that for polar transformations the angular derivative is multiplied by the radius.

The hue, h, is a well known full shadow-shading-specular invariant. Eq. 3.15 provides a link between the derivative of the full invariant, h_x and the quasi-invariant $|\mathbf{H}_x^c|$. A drawback of hue is its undefinedness for points on the black-white axis, i.e. for small s. Therefore the derivative of hue is unbounded. In section 3.3, we derived the quasi-invariant as a linear projection of the spatial derivative. For these projections, it holds that $0 < |\mathbf{H}_x^c| < |\mathbf{f}_x|$, and hence the shadow-shading specular quasi-invariant is bounded. It should be mentioned that small changes round the grey axis, result in large changes of the direction or 'color' of the derivative, e.g. from blue to red, in both the quasi-invariant and the full invariant. However, the advantage of the quasi-invariant is that the norm remains bounded for these cases. For example, in
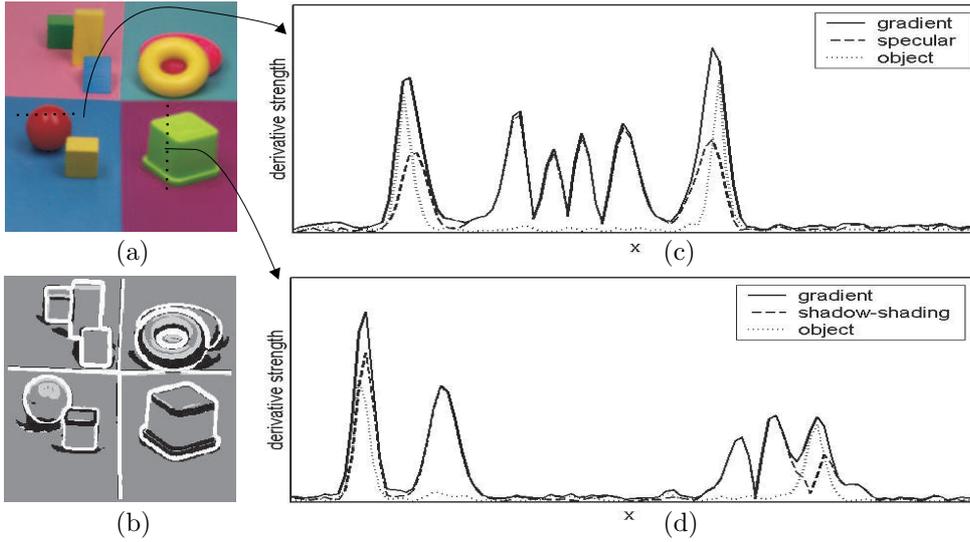
Figure 3.4: *(a) Input image with superimposed two dotted lines which are plotted in the images (c) and (d). (b) Edge classification result, with white object edges, black shadow edges and light grey specular edges. (c),(d) The derivative strength along lines indicated in (a) (see also color plate C.6).*

Fig. 3.3d a red-blue edge is depicted. The blue patch becomes more achromatic along the y-axis. The instability for grey values is clearly visible in Fig. 3.3e whereas in Fig. 3.3f the response of the quasi-invariant remains stable.

3.4.4   Characteristics of Quasi-Invariants

Full invariants are invariant with respect to a physical photometric parameter like for instance the geometric term m^b in the case of normalized RGB. Hence, the first order derivative response of such invariants does not contain any shadow-shading variation. Our approach determines the direction in the RGB-cube in which shadow-shading edges exhibit themselves. This direction is then used to compute the quasi-derivative which shares with full invariants the property that shadow-shading edges are ignored. However, the quasi-invariants is not invariant with respect to m_b. For the shadow-shading quasi-invariant subtraction from Eq. 3.3 of the part in the shadow-shading direction $\mathbf{c}^b$ results in

$$\mathbf{f}_x = em^b \left(\mathbf{c}_x^b - \mathbf{c}_x^b \cdot \hat{\mathbf{c}}^b \right) \tag{3.16}$$

which is clearly not invariant for m^b and e. In a similar way also the specular-shadow-shading quasi-invariant can be proven to be dependent on m^b and e.

The dependency of the quasi-invariants on m^b and e limits their applicability. They cannot be used for applications where edge responses are compared under different circumstances, such as content based image retrieval. However, they can be used in applications which are based on a single frame, such as shadow-edge insensitive

image segmentation, shadow-shading-specular independent corner detection and edge classification.

A major advantage of the quasi-invariants is that their response to noise is independent of the signal. In the case of additive uniform noise, the noise in the quasi invariants is also additive and uniform, since it is a linear projection of the derivative of the image. This means that the noise distortion is constant over the image. In section 3.4, it was shown that the full invariants differ from the quasi-invariants by scaling with a signal depended factor (the intensity or saturation). And hence their noise response is also signal depended. Typically the shadow-shading full invariant exhibits high noise distortion round low intensities while the shadow-shading-specular full invariant has high noise dependency for points around the achromatic axis. This is shown in Fig. 3.3. The uneven levels of noise throughout an image hinder further processing.

A second advantage of photometric variants and quasi-invariants is that they are expressed in the same units (i.e. being projections of the derivative they are in RGB-value per pixel). This allows for a quantitative comparison of their responses. An example is given in Fig. 3.4. Responses along two lines in the image are enlarged in Fig. 3.4c and Fig. 3.4d. The line in Fig. 3.4c crosses two object edges and several specular edges. It nicely shows that the specular-variant almost perfectly follows the total derivative energy for the specular edges in the middle of the line. In Fig. 3.4d a line is depicted which crosses two object edges and three shadow-shading edges. Again the shadow-shading variant follows the gradient for the three shading edges. A simple classification scheme results in Fig. 3.4b. Note that full-invariants cannot be compared quantitatively because they have different units.

3.5   Experiments

We compare the performance of the quasi-invariants with the full invariants according to the following criteria 1. stability, 2. edge displacement and 3. discriminative power. For the improved stability a mathematical proof is given in chapter 3.4. Here, we will test the invariants on edge displacement and discriminative power.

Since the specular quasi-invariant is well-known, and it does not counterpart a full invariant, its performance is not investigated here. The experiments were performed with normalized RGB, $c_1 c_2 c_3$, $l_1 l_2 l_3$, hue, Cw and Hw [16] [18]. The results for the invariants $c_1 c_2 c_3$, $l_1 l_2 l_3$, Cw and Hw were similar or worse than the results for normalized RGB and hue. Therefore, we have chosen normalized RGB and hue as exemplary for the set of invariants, and compared them with the quasi-invariants. Implementation details of the quasi-invariants can be found in [74]. For the experiments a white light source $\hat{\mathbf{c}}^i = \frac{1}{\sqrt{3}}(1, 1, 1)^T$ is used.

3.5.1   Edge Detection

First, we compare the edge detection performance of the quasi-invariants with the invariants from literature. These results can also be seen as an indication of the

std. noise →	5		20	
detector ↓	Δ	ε	Δ	ε
S_x^c	0.043	0.99%	0.43	10%
rgb	0.21	2.0%	1.1	18%
H_x^c	0.35	5.8%	0.98	20%
hue	0.85	9.8%	2.1	34%
$RGB - gradient$	0.003	.07%	0.08	2.0%

Table 3.1: *The displacement, Δ, and the percentage of missed edges, ε, for five different edge detectors. Gaussian noise of standard deviation 5, and 20 was added.*

loss of discrimination due to invariance. Edge detection is performed between the 1012 different colors from the PANTONE [58] color system. Every one of the 1012 different RGB-values is combined with all other RGB-values, resulting in a total of $N = 1012 * 1011/2 = 511566$ edges of $M = 25$ pixels length. The edge position is determined by computing the maximum response path of the derivative energy in a region of 20 pixels around the actual edge. This results in an edge estimation which is compared with the actual edge. We define two error measures. First, the average pixel displacement Δ,

$$\Delta = \frac{\sum\limits_{\{x_{i,j}; |x_{i,j} - x_0| > 0.5\}} |x_{i,j} - x_0|}{N \cdot M} \tag{3.17}$$

in which $x_{i,j}$ is j-th edge pixel of the i-th edge. Because the actual edge is located between two pixels displacements equal to .5 pixels are considered as a perfect match. The second error measure is the percentage of missed edges, ε. An edge was classified missed as the variation over one edge,

$$\text{var}(i) = \frac{1}{M} \sum_{j=1}^{M} \left| x_{i,j} - \frac{1}{M} \sum_k x_{i,k} \right| \tag{3.18}$$

is larger than 1 pixel. For the Gaussian derivative, a scale $\sigma = 1$ is chosen. The experiments were performed with uncorrelated Gaussian noise of standard deviation 5, and 20.

The results are depicted in Table 3.1. For both cases, the shadow-shading and shadow-shading-specular edges, the quasi-invariants substantially outperform the invariants. For comparison, the results without invariance based on the RGB gradient, $|\mathbf{f}_x|$, are inserted. Obviously, the RGB gradient has the best discriminative power. However, it will also find many edges which are caused due to scene incidental events.

To provide more insight in what kind of edges were still detected, we computed the average Euclidean RGB difference of the missed edges for the case with Gaussian noise with a standard deviation of 5. With $d_{ij} = |\mathbf{f}_i - \mathbf{f}_j|$ the Euclidean distance between patch i and j. For the RGB gradient-based method, we obtained an average distance of $\overline{d} = 4.6$, for the shadow-shading quasi-invariant $\overline{d} = 86$ and $\overline{d} = 109$ for the shadow-shading-specular invariant.

| (a) | (b) | (c) | (d) |
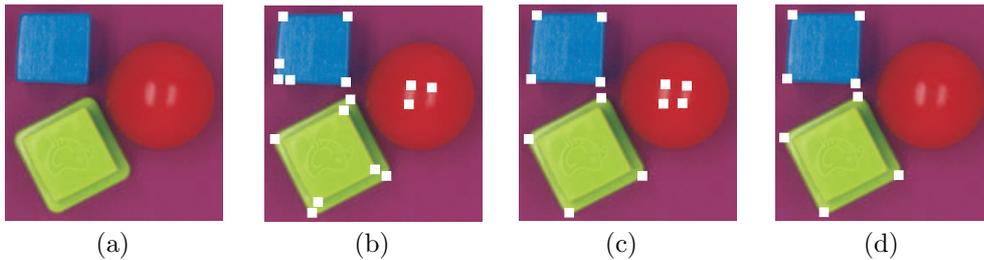
Figure 3.5: *(a) Input image and corner detector results based on (b) RGB gradient* $(\mathbf{f}_x)$, *(c) shadow-shading quasi-invariant* $(\mathbf{S}_x^c)$, *and (d) shadow-shading-specular quasi-invariant* $(\mathbf{H}_x^c)$ *(see also color plate C.7).*

3.5.2   Photometric invariant corner detection

Derivatives based on full photometric invariants are, due to their instability, unreliable input for geometrical operations such as photometric invariant corner detection, orientation estimation, curvature estimation, etc. Quasi-invariants, on the other hand are expected to be more stable in combination with geometrical operations. We used the following straightforward extension of the Harris corner detector [27] for color images

$$H\mathbf{f} = \overline{\mathbf{f}_x^T \mathbf{f}_x}\; \overline{\mathbf{f}_y^T \mathbf{f}_y} - \overline{\mathbf{f}_x^T \mathbf{f}_y}^2 - k(\overline{\mathbf{f}_x^T \mathbf{f}_x} + \overline{\mathbf{f}_y^T \mathbf{f}_y})^2. \tag{3.19}$$

The overline indicates a gaussian averaging window. The corner detection results are given in Fig. 3.5. The shadow-shading quasi-invariant detector does not find shadow-shading corners whereas the shadow-shading-specular quasi-invariant also ignores the specular corners.

In Fig. 3.6, the 30 most prominent Harris corners are detected for two real world images (Corel gallery). The detected points can be used as interest points for object recognition [64]. Note that the images break several of the assumptions of the dichromatic reflection model (Eq. 3.1). They do not have a known illuminant, nor are they taken with a linear acquisition system. The results for the full invariants are dominated by their instabilities. The shadow-shading full invariant is unstable in the low intensity areas and consequently finds most of the interest points in this area. The shadow-shading specular full invariant is unstable along the whole grey axis, which leads to false corners in grey areas. The *RGB* gradient method focusses on large *RGB* value changes which mostly coincide with light-dark transition which are rarely the most discriminative points. It is apparent that the quasi-invariants (Figs. 3.6d,h) suppress unwanted photometric variation and focus on body reflectance changes only.

3.6   Conclusions

In this chapter we proposed a set of quasi-invariant derivatives. These derivative filters are combined with derivative-based feature detectors to perform photometric
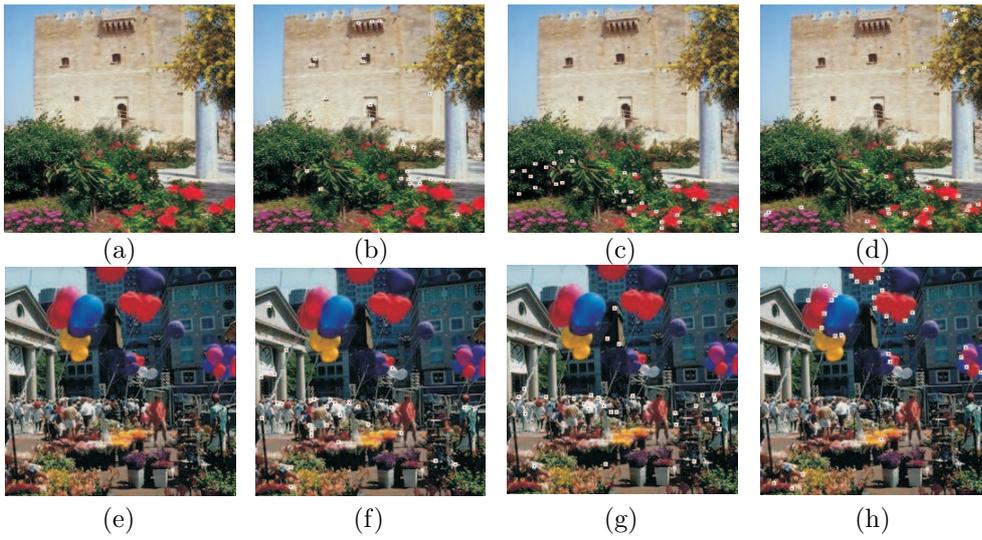
Figure 3.6: *(a), (e) Input images. Corner detection based on (b) RGB gradient ($\mathbf{f}_x$), (c) normalized RGB, (d) shadow-shading quasi-invariant ($\mathbf{S}_x^c$), (f) RGB gradient ($\mathbf{f}_x$), (g) hue full invariant (h_x), and (h) shadow-shading quasi-invariant ($\mathbf{H}_x^c$) (see also color plate C.8).*

invariant feature detection. Experiments show that they significantly outperform feature detection based on full invariants on both stability and discriminative power.

Chapter 4

Curvature Estimation in Oriented Patterns using Curvilinear Models applied to Gradient Vector Fields *

4.1 Introduction

Reliable estimation of local features in digitized images is of great importance for many image processing tasks (segmentation, analysis, and classification). Depending on the class of images under investigation, knowledge of different features is desired. One such class of images is defined by Kass and Witkin [37] as oriented patterns: patterns that exhibit a dominant local orientation. Examples are seismic, acoustic, wood grain, interference patterns, and fingerprint images. Important features for these images are estimates of local anisotropy, orientation, curvature and scale.

The structure tensor yields a robust estimator for local orientation [7] [23] [37] [87] based on a local gradient vector field. This estimator locally models the images as translation invariant strokes. In addition to orientation estimation this method also yields an anisotropy measure indicating the resemblance of the local area to a translation invariant model. This measure can also be interpreted as a confidence measure of the estimated orientation. Both a lack of smoothness (e.g. caused by noise) and deviations from the translation invariant model (e.g. curved oriented patterns) are responsible for a decrease of this confidence measure. To distinguish between the two

---

possible causes, we proposed a parabolic transformation, which optimizes the translation invariance after transformation [90]. This method yields a curvature estimate for curved oriented patterns as a by-product. A shortcoming of this method is that the proposed transformation is not orthonormal and therefore it lacks conservation of gradient energy. This does not allow direct comparison of the confidence values of different transformations. In this paper we propose a method to investigate the resemblance of a local pattern of 2-D oriented pattern to a certain model function (e.g. circular, parabolic). The model is represented by a parameterized transformation function of the isophotes. The method assures the conservation of gradient energy, allowing us to compare confidence measures of different transformations, and especially of a parameterized transformation for different parameter values. Like in [90], solving the parameter for best confidence yields a closed-form estimate of the additional free parameter, e.g. local curvature. We propose two curvilinear models, a parabolic and a circular model, for the characterization of curved oriented patterns. When the resemblance between a model and a local image is high, the corresponding model parameters, orientation and curvature, yield a reliable description of the local image. The method yields features with a corresponding confidence value. All these estimates are local and can be represented as feature maps.

Estimation of the curvature in oriented patterns is not trivial. Worring [94] presented an extensive comparison between curvature estimators applied to segmented data for which the position and ordering of points along the contour have to be known. For noisy oriented patterns segmentation fails, making these methods useless. The isophotes (tangential) curvature (the second derivative along the isophotes divided by the gradient magnitude) is segmentation-free [86][89], but also fails on these images. There are three reasons for this [90]: a) the gradient is zero on ridges and in valleys; b) Increasing the regularization scale of directional derivatives suppresses the oriented pattern and reduces the signal-to-noise ratio; c) opposite sides of a ridge (or valley) yield curvatures of opposite sign, which cancel out after averaging. The only two methods which do yield a curvature estimate for oriented patterns are either very computationally demanding [83] or are not accompanied by a confidence measure, which makes them hard to rely on [88].

The proposed method resembles a method for the detection of complex symmetries as presented by Bign [6][7][24]. He characterizes symmetries by (coordinate-) transformation functions, which transform symmetric patterns into translation invariant patterns. The success of such a transformation is determined by the confidence measure of the structure tensor applied to the transformed image. A high confidence value is an indicator for the presence of the corresponding symmetry. Bign's method is an extension of the generalized Hough transform. Detection of a symmetry pattern involves accumulation of evidence by voting. Bign's symmetry detector requires two orthonormal transformation functions. It measures the resemblance of the local differential field to two perpendicular differential fields (indicating the symmetry), whereas our method looks at the resemblance of the local differential field to only one differential field (representing the shape of the isophotes). This difference allows us to estimate model parameters by optimizing the resemblance between the actual differential field and a model differential field in a closed-form solution, i.e., omitting

a time consuming voting scheme. This is not possible with the symmetry method since neither one of the two differential fields is preferred. The requirement for two orthonormal transformation functions poses an unnecessary limitation to the symmetries. For example, such a set of functions does not exist for the parabolic model we propose, i.e. parabolic isophotes along a linearly increasing symmetry axis. We extend his method by noting that only the existence of the differential fields of the two transformation functions is essential.

4.2   Oriented Patterns

A oriented pattern $m(x, y)$ can be written as a real one dimensional function g of a model function u

$$m(x, y) = g(u(x, y, \mathbf{a})).\qquad(4.1)$$

The model function $u(x, y, \mathbf{a})$ describes the shape of the isophotes and $\mathbf{a}$ contains local isophotes parameters such as orientation and curvature. Consequently, the gradient (differential field) of m

$$\nabla m = \frac{dg}{du}\,\nabla u \qquad(4.2)$$

is a $\frac{dg}{du}$ weighted version of the gradient of u. In oriented patterns we distinguish between two perpendicular orientations; along the isophotes (tangent), and along the gradient. Note that orientation is defined on the interval $[0, \pi\rangle$. Consequently, vectors in opposite directions have the same orientation.

   Consider the function $f(x, y)$ representing a local image (window) and a model function $u(x, y, \mathbf{a})$. It is of interest to what extent $f(x, y)$ is described by the model function $u(x, y, \mathbf{a})$. This is measured by decomposing the derivative energy of $f(x, y)$ into two contributions, one parallel and one perpendicular to the normalized differential field of $u(x, y, \mathbf{a})$. This results in the following energies

$$\begin{aligned}
E_f(\mathbf{a}) &= \int\int \left(\nabla f \cdot \frac{\nabla u(\mathbf{a})}{\|\nabla u(\mathbf{a})\|}\right)^2 dx\, dy \\
E_r(\mathbf{a}) &= \int\int \left(\nabla f \cdot \frac{(\nabla u(\mathbf{a}))_\perp}{\|\nabla u(\mathbf{a})\|}\right)^2 dx\, dy
\end{aligned}\qquad(4.3)$$

where $E_f(\mathbf{a})$ denotes the fit energy and $E_r(\mathbf{a})$ the residual energy. The subscript $\perp$ indicates a rotation of 90∘ of the vector and the integrals represent the averaging over the local image. A quality measure of the fit can be found by comparing the fit energy with the residual energy. Since no a-priori knowledge exists to interpret the energy difference between the fit and the residual energy, we normalize the difference with the total gradient energy to obtain the following quality measure $c(\mathbf{a})$

$$c(\mathbf{a}) = \frac{E_f(\mathbf{a}) - E_r(\mathbf{a})}{E_f(\mathbf{a}) + E_r(\mathbf{a})}\qquad -1 \leq c \leq 1.\qquad(4.4)$$

The value of $c(\mathbf{a})$ varies from -1 for a pattern of which the isophotes are exactly perpendicular to those of the model function $u(x, y, \mathbf{a})$ and $+1$ for a pattern which

is exactly described by the model function. The isotropic noise energy is distributed equally between the fit and the residual energy.

More important than the quality measure for an arbitrary $\mathbf{a}$ is to know which $\mathbf{a}$ maximizes the quality function c, i.e. maximizes E_f and minimizes E_r. The vector $\mathbf{a}$ contains model parameters which describe local features. Therefore optimizing the confidence function c corresponds to feature estimation. Furthermore, the quality measure $c(\mathbf{a})$ informs us about the success of the fit and can be seen as a confidence measure of the estimated features. Besides comparing confidence measures of the same model function, it is also possible to compare confidence measures of different model functions. Note that the normalization of the confidence measures is independent of the model function. By comparing optimized confidence functions of various models one can find out which model describes the local pattern best.

Usually the complexity of the confidence function does not allow a closed-form solution of the optimization criterion. The straight model is an exception. In the case of curvilinear models, we avoid costly (iterative) optimization procedures by considering approximate confidence functions, which do allow closed-form solutions.

4.3   Straight-Oriented Patterns

Locally, many oriented patterns can be characterized by a straight model. For such a pattern the model function $u(x, y, \mathbf{a})$ is given by

$$u(x, y, \phi) = x \cos \phi + y \sin \phi \tag{4.5}$$

with ϕ the orientation perpendicular to the model isophotes. Substituting this in Eq. (3) yields

$$E_f(\phi) = \tfrac{1}{2}\left(\overline{f_x^2} + \overline{f_y^2}\right) + \tfrac{1}{2}\left(\overline{f_x^2} - \overline{f_y^2}\right)\cos 2\phi + \tfrac{1}{2}\overline{2f_x f_y}\sin 2\phi. \tag{4.6}$$

A bar $\overline{(.)}$ denotes an averaged quantity and will from now on replace the integrals responsible for averaging over a local image. The confidence value $c(\phi)$ is

$$c(\phi) = \frac{1}{\overline{f_x^2} + \overline{f_y^2}}\left(\left(\overline{f_x^2} - \overline{f_y^2}\right)\cos 2\phi + \overline{2f_x f_y}\sin 2\phi\right) \tag{4.7}$$

$c(\phi)$ can be maximized as a function of the orientation ϕ. This yields the following (gradient-based) orientation estimator[8][23] [37][87]

$$\phi_{opt} = \tfrac{1}{2}\arctan\frac{2\overline{f_x f_y}}{\overline{f_x^2} - \overline{f_y^2}} \tag{4.8}$$

with confidence value $c(\phi_{opt})$

$$c(\phi_{opt}) = \frac{d^2}{g^2} \quad where \quad d^4 = \overline{f_x^2 - f_y^2}^2 + \overline{2f_x f_y}^2. \tag{4.9}$$

This confidence measure can also be interpreted as a measure for translation invariance and shows an intuitive dependency to the pattern orientation ϕ_{opt}.

$$c\left(\phi\right) = \frac{d^2\left(\cos^2\left(\phi - \phi_{opt}\right) - \sin^2\left(\phi - \phi_{opt}\right)\right)}{g^2} = \tfrac{1}{2}c\left(\phi_{opt}\right)\left(1 + \cos\left(2\left(\phi - \phi_{opt}\right)\right)\right).$$
(4.10)

The maximum of the confidence measure $c\left(\phi_{opt}\right)$ reduces due to noise in the local image f. For a linear pattern p distorted by additive uncorrelated noise n ($f = p + n$) the confidence value $c\left(\phi_{opt}\right)$ is

$$c\left(\phi_{opt}\right) = \frac{d^2}{\left\|\nabla f\right\|^2} = \frac{d^2}{\left\|\nabla p + \nabla n\right\|^2} = \frac{d^2}{\left\|\nabla p\right\|^2}\frac{\overline{\left\|\nabla p\right\|^2}}{\overline{\left\|\nabla p\right\|^2} + \overline{\left\|\nabla n\right\|^2}}.$$
(4.11)

Note that the gradient noise energy is divided equally over E_f and E_r. Therefore the numerator of c is unaffected by noise. Noise increases the total gradient energy (denominator of c), which lowers the confidence value $c\left(\phi_{opt}\right)$. Another reason for a lower confidence value is a deviation between the local image and the model function. For instance when curved lines occur, then curvature will contribute to E_r. In the next section we will extend the model to include curved patterns.

4.4   Curved Oriented Patterns

We present two model functions, which locally model curved oriented patterns. A parabolic model

$$u\left(x, y, \phi, \kappa\right) = \tfrac{1}{2}\kappa w^2 - v$$
(4.12)

and a concentric circle model

$$u\left(x, y, \phi, \kappa\right) = \sqrt{\kappa^2 w^2 + \left(1 - \kappa v\right)^2}$$
(4.13)

in which κ is the curvature. The Gauge coordinates v, w are obtained by

$$v = x\cos\phi + y\sin\phi \qquad w = -x\sin\phi + y\cos\phi.$$
(4.14)

Here we discuss the parabolic approximation. For the circular approximation we refer to appendix A. Using the parabolic model function and Eq. (3) the following energies are obtained

$$\begin{aligned}
E_f\left(\phi, \kappa\right) &= \overline{\left(\frac{\kappa^2 w^2 f_w^2 - 2\kappa w f_w f_v + f_v^2}{1 + \kappa^2 w^2}\right)} \\
E_r\left(\phi, \kappa\right) &= \overline{\left(\frac{\kappa^2 w^2 f_v^2 + 2\kappa w f_v f_w + f_w^2}{1 + \kappa^2 w^2}\right)}
\end{aligned}$$
(4.15)

where f_v and f_w are the derivatives in respectively the v and w direction. Finding the curvature and orientation that maximize the confidence function requires a search in ϕ, κ-space. In this paper we shall not further investigate this method due to its high computational demands. Instead we propose a way to approximate the confidence function, allowing a fast closed-form solution.

An approximation to the orientation ϕ can be obtained by looking at the axis of minimal translation invariance for parabolic and circular patterns. In the case of a circular pattern this is the v-axis. For a parabolic pattern it depends on the curvature and the window size used. For small curvatures (i.e. compared to the window size) the minimal translation invariant axis is equal to the v-axis. Increasing the curvature the axis of minimal translation invariance jumps to the w-axis. Therefore an approximation of the orientation needed to determine the v and w-axes in the Eq. (15) can be computed with Eq. (8). After substituting the orientation, the resulting equations only depend on the curvature. Iterative maximization of the confidence function in κ-space is still time-consuming. We propose to approximate this maximum by using locally adapted weighting. The weighting function of E_f and E_r (denoted by the bar $\overline{(.)}$) is on its turn weighted by $\left(1 + \kappa^2 w^2\right)$ after which we normalize for this weighting. This mathematical trick has a high resemblance to normalized convolution [39]. It results in

$$
\begin{aligned}
\hat{E}_f(\kappa) &= \frac{\kappa^2 \overline{w^2 f_w^2} - 2\kappa \overline{w f_w f_v} + \overline{f_v^2}}{1 + \kappa^2 \overline{w^2}} \\
\hat{E}_r(\kappa) &= \frac{\kappa^2 \overline{w^2 f_v^2} + 2\kappa \overline{w f_w f_v} + \overline{f_w^2}}{1 + \kappa^2 \overline{w^2}}.
\end{aligned}
\tag{4.16}
$$

A hat ($\hat{}$) above a quantity indicates an approximation. Since the fit energy E_f might be a function of the coordinate w, as is the adapted weighting function, optimization lead to a false curvature estimate. Therefore minimization of the residual energy E_r is used to find the following closed-form curvature estimate

$$
\hat{\kappa} = \frac{\overline{w^2 f_v^2} - \overline{w^2} \cdot \overline{f_w^2} - \sqrt{4\overline{w^2} \cdot \overline{w f_w f_v}^2 + \left(-\overline{w^2 f_v^2} + \overline{w^2} \cdot \overline{f_w^2}\right)^2}}{2\overline{w^2} \cdot \overline{w f_w f_v}}.
\tag{4.17}
$$

The confidence measure can now be computed in two different ways. The confidence measure $c(\phi, \kappa)$ has its maximum at $(\phi_{opt}, \kappa_{opt})$. To avoid an iterative search for this optimum one can compute $c(\hat{\phi}, \hat{\kappa})$ by substituting $\hat{\phi}$ and $\hat{\kappa}$ in Eqs (15) and (4). Note that estimates $\hat{\phi}$ and $\hat{\kappa}$ do not have to be equal to the values of ϕ and κ that optimize the confidence function. However, computing $c(\hat{\phi}, \hat{\kappa})$ is still expensive. A significant speed-up can be obtained by approximating the confidence measure using the approximate energies of Eq. (16).

$$
\hat{c}(\phi, \kappa) = \frac{\hat{E}_f(\phi, \kappa) - \hat{E}_r(\phi, \kappa)}{\hat{E}_f(\phi, \kappa) + \hat{E}_r(\phi, \kappa)}.
\tag{4.18}
$$

Again, one can avoid an iterative search by substituting $\hat{\phi}$ and $\hat{\kappa}$ in Eq. (18), which yield $\hat{c}(\hat{\phi}, \hat{\kappa})$. The curvature estimator in Eq. (17) is the tangential or isophote curvature. The normal (or gradient flow line) curvature can be computed by exchanging the v and w coordinates in Eqs (12) and (13).

4.5 Implementation

Direct computation of the curvature and the confidence measure is a space-variant operation. This yields a high computational demand. Fortunately, Eqs (16) to (18) can be calculated with global convolutions, which can be implemented by multiplication in the Fourier-domain. This yields a substantial reduction in computational complexity. The derivatives f_x and f_y are implemented as regularized derivative filters.

$$f_x \equiv f(x,y) \otimes \frac{\partial g(x,y;\sigma_g)}{\partial x} \overset{F}{\leftrightarrow} j\omega_x \tilde{f}(\omega_x,\omega_y)\, \tilde{g}(\omega_x,\omega_y;\sigma_g) \qquad (4.19)$$

with $\tilde{f}$ the Fourier transform of f and $g(x,y;\sigma_g)$ a Gaussian regularization function of scale σ_g

$$g(x,y;\sigma_g) = \frac{1}{2\pi\sigma_g^2} e^{-(x^2+y^2)/2\sigma_g^2} \overset{F}{\leftrightarrow} \tilde{g}(\omega_x,\omega_y;\sigma_g) = e^{-\frac{1}{2}\left(\omega_x^2+\omega_y^2\right)\sigma_g^2}. \qquad (4.20)$$

The terms of the curvature estimator and the confidence measure, Eqs (16)-(17) , are expanded in Appendix B (the circular model is treated in appendix A). The remaining terms $\overline{x^p y^q f_x^r f_y^s}$ are conveniently calculated as multiplications in the Fourier domain

$$\overline{x^p y^q f_x^r f_y^s} = u(p,q,\sigma_a) \otimes f_x^r f_y^s \overset{F}{\leftrightarrow} \tilde{u}(p,q,\sigma_a) F\left\{f_x^r f_y^s\right\}. \qquad (4.21)$$

For the window function we choose a Gaussian of scale σ_a. The Fourier transform of the filter u is

$$u(p,q,\sigma_a) \equiv x^p y^q g(x,y;\sigma_a) \overset{F}{\leftrightarrow} \tilde{u}(p,q;\sigma_t) \equiv j^{p+q} \frac{\partial^{p+q}\tilde{g}(\omega_x,\omega_y;\sigma_a)}{\partial\omega_x^p \partial\omega_\eta^q}. \qquad (4.22)$$

Due to the high frequency character of oriented patterns σ_g should be kept small, i.e. tuned to the frequency characteristics of the cross-section of a line. Noise suppression is accomplished by averaging all terms by Gaussian window (size σ_a), i.e. the size of the curvilinear model.

4.6 Experiments

In this section the proposed algorithms are tested on synthetic and real-world images. The feature extraction, which we presented, is based upon finding a maximum of the confidence measure in parameter space $c(\mathbf{a})$. The curvature of oriented patterns corresponds to the position of the maximum in $c(\kappa,\phi)$. To avoid searching κ,ϕ-space the approximations $\hat{\phi}$ and $\hat{\kappa}$ are proposed. With these an approximated confidence measure $\hat{c}$ and the exact confidence measure c may be computed. The goal of the experiments is to investigate the performance of these approximations as a function of the curvature. Also the robustness with respect to the noise is checked. The tests are performed on a concentric circle image $f(x,y) = \sin\left(\sqrt{x^2+y^2}+\varphi\right)+n$ (see Fig. 1) in which $n = N\left(0,\sigma_n^2\right)$ and φ a phase-term set randomly for every noise-realization.
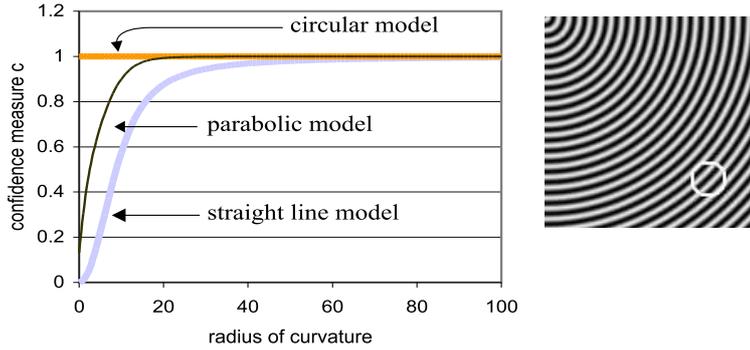
Figure 4.1: *Confidence measure $c(\hat{\phi}, \hat{\kappa})$ of circular, parabolic, and straight line models on a noise-free pattern of concentric circles.*

For the signal-to-noise ratio we use $SNR = 10 \log \left(h^2 / \sigma_n^2 \right)$ where h is the contrast difference and σ_n the standard deviation of the noise. Be aware that the proposed algorithms are based on the gradient energy of the local image. Thus an increase of the pattern frequency will usually result in a higher SNR (gradient energy vs. filtered noise variance) and therefore a better performance. All experiments on the concentric circle image are based on 100 measurements. Unless mentioned otherwise the sigma sizes are $\sigma_g = 1.0$ and $\sigma_a = 5.0$.

4.6.1   Confidence measure as selection criterion

The importance of choosing the right model is illustrated in Fig. 1, which shows the confidence measures of the circular, parabolic and the straight model applied to a noise-free pattern of concentric circles. It is clear that for high curvatures the deviation of the straight and the parabolic model form the circle pattern results in a significantly lower value of the confidence measure.

4.6.2   Bias of the Actual Confidence Measure

To investigate to what extent the optimum of the confidence function in κ, ϕ-space is found, we compare the average confidence measure of the circular model applied to curved patterns with the average confidence measure of a straight model applied to straight pattern. Both images have identical signal-to-noise ratios. The confidence measure $c(\phi, \kappa)$ of a curvilinear model can be slightly higher than the confidence measure of a straight model. This slight increase is caused by the fact that the curved model allows for two parameters to adjust to the noise.

The average confidence measure $c(\hat{\phi}, \hat{\kappa})$ of the circular model applied to the concentric circles is depicted in Fig. 2 for three SNR's (20dB, 10dB, 6dB). It clearly shows that for small radii the average confidence measure $c(\hat{\phi}, \hat{\kappa})$ decreases. This is

caused by an increasing discrepancy between the approximated $(\hat{\phi}, \hat{\kappa})$ and the optimal $(\phi_{opt}, \kappa_{opt})$ for small radii. Note, $c(\phi_{opt}, \kappa_{opt})$ does not decrease for small radii. Fig. 2b indicates the variation around the average confidence measure for the straight model. Increasing the window size (local image) reduces the variation in exchange of a further decrease of $c(\hat{\phi}, \hat{\kappa})$ for small radii.

4.6.3 Approximation Error of the Confidence Measure

In section 4 we presented two methods for computing the confidence measure, the actual confidence measure $c(\hat{\phi}, \hat{\kappa})$ and an approximation $\hat{c}(\hat{\phi}, \hat{\kappa})$. In Fig. 3 the rms (root-mean-square) error due to this approximation is depicted for the circular and the parabolic model. For both models these errors are small. Only for high curvatures (small radii) it may be worthwhile to compute the actual confidence measure.

4.6.4 Robustness of the Curvature Estimator

It is important to test the robustness of the curvature estimation. In Fig. 4, the noise sensitivity of the parabolic and circular curvature estimators are depicted. Both models were applied to the concentric circles. The coefficient-of-variation $(CV = \sigma/\mu)$ of both models are similar for the middle and high SNR's, but the parabolic models performs better for low SNR's. Considering the advantage of the circular curvature estimator due to the exact match between the model and the pattern, we show that parabolic curvature estimator suffers less from the approximations. The parabolic
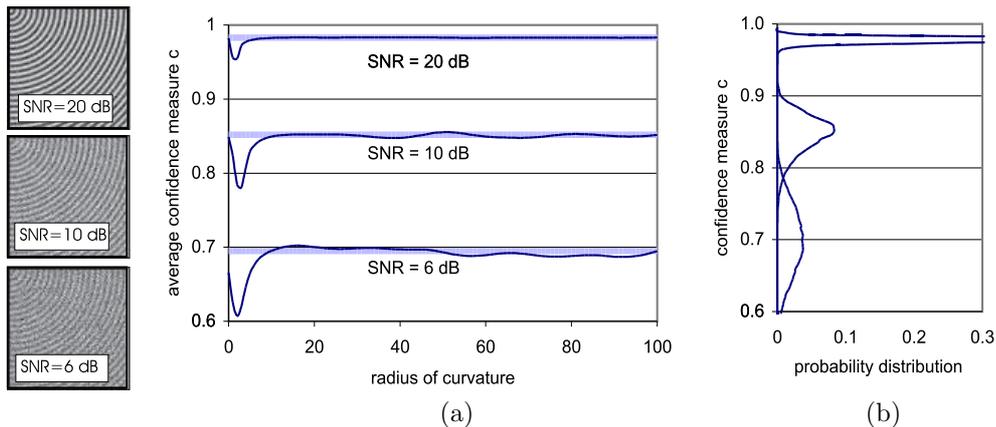
Figure 4.2: *(a) Average confidence measure $c(\hat{\phi}, \hat{\kappa})$ for the circular model as a function of the radius for three different SNRs (top to bottom: 20 dB, 10 dB, 6 dB). The measure $c(\hat{\phi}, \hat{\kappa})$ yields a small bias for small radii. The horizontal lines indicate the average confidence measure for the straight-line model for the corresponding SNR. (b) Probability density functions of the confidence measures for the straight-oriented patterns for the three different SNRs (top to bottom: 20 dB, 10 dB, 6 dB).*
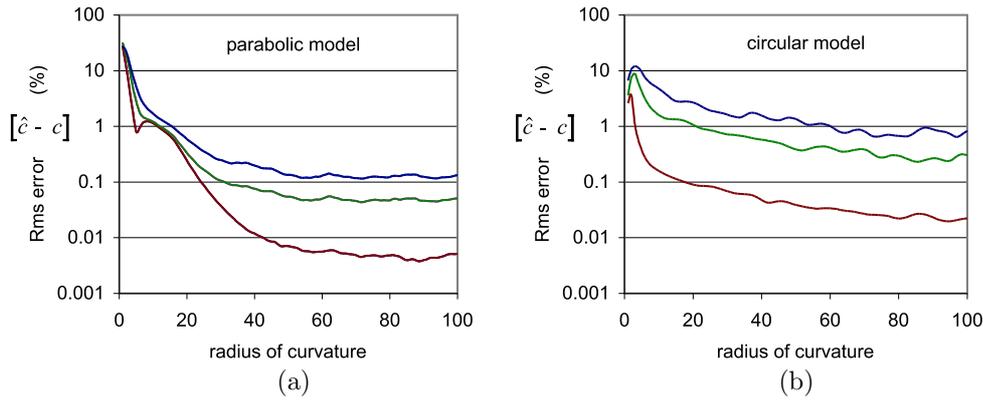
Figure 4.3: *Rms error between the actual confidence measure and its approximation as a functioin of the radius. (a) Approximate parabolic model applied to concentric circle pattern of various SNR (top-to-bottom: 6dB, 10 dB, 20 dB). (b) Approximate circular model applied to concentric circle patterns of various SNR (top-to-bottom: 6dB, 10 dB, 20 dB).*

curvature estimator performs at least as well over a wide range of curvatures. Only for high curvatures the circular model can take advantage of the exact match. In practice, one can compute the curvature corresponding to both models. The one with the highest confidence measure is preferred because its model yields a better description of the data.

4.6.5   Application of Curvilinear Models to Real-World Data Sets

In Fig. 5 an interference pattern, together with the curvature and confidence estimation for both the parabolic and circular model, is depicted. As expected, the parabolic model fails in the middle of the ellipses as indicated by an abrupt drop of the confidence measure. The circular confidence measure hardly decreases for the circles at the top and the bottom of the image. For the flatter ellipses on the left and the right the mismatch between the model and the pattern is slightly larger. In the difference image between the circular and parabolic confidence measures, the lighter areas indicate a better description of the circular model whereas in the darker areas the parabolic model yields a better fit. The slightly darker lines denote an almost perfect parabolic line pattern.

The estimated local curvature of a fingerprint and a CT cross-section of a tree-trunk are depicted in Fig. 6. Both curvilinear models produced similar results. The dark lines in the logarithmically stretched curvature images denote locally straight patterns. Both peaks in the fingerprints curvature correspond to important minutia for fingerprint recognition [35] [49]. The curvature estimation can be used to improve (to prevent jumping the rails) the ridge tracking [35], which is already based on orientation estimation. The high confidence measures (white areas in confidence images)
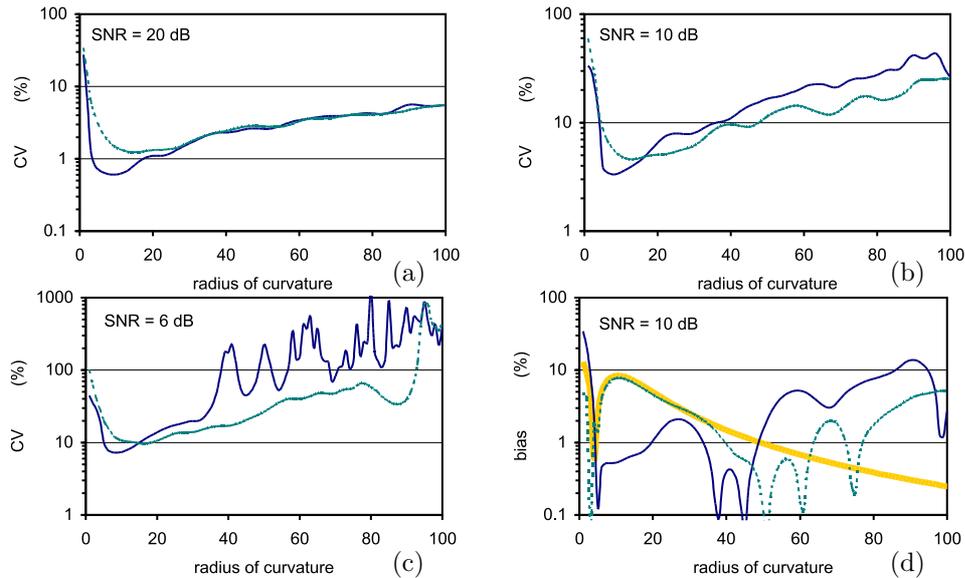
Figure 4.4: *Curvature estimators using curvilinear models: black line = parabolic model, gray line = circular model. (a), (b) and (c) Coefficicient-of-variation (CV) for the parabolic and circular model-based curvature estimators for different SNR (20 dB, 10 dB, 6 dB). (d) Bias of parabolic and circular model based curvature estimators (SNR = 10 dB) (thick gray line indicates the noise-free bias of curvature using the parabolic model).*

indicate a perfect fit of the model and a reliable curvature estimate.

4.7   Conclusions

In this paper we present a method to compare a local image with a model function. A quality measure indicates the resemblance between the local image and the model function. Feature extraction is obtained by optimization of the quality function as a function of the parameters, which represent the feature. The quality function is interpreted as a confidence measure for the measured features. We propose two curvilinear models to describe curved oriented patterns. To avoid searching ϕ, κ-space we propose closed-form solution for approximations to the actual parameters of the curvilinear models $\hat{\phi}$ and $\hat{\kappa}$. Instead of the exact confidence measure $c(\hat{\phi}, \hat{\kappa})$ an approximation $\hat{c}(\hat{\phi}, \hat{\kappa})$ can be computed resulting in a huge reduction in computational demand. We demonstrate that these approximations yield good results for almost all curvatures. Only for the highest curvatures one might decide to compute $c(\hat{\phi}, \hat{\kappa})$, or (even more computationally demanding) to iterate in ϕ, κ-space for $c(\phi_{opt}, \kappa_{opt})$.

(a)　　　　　　　　(b)　　　　　　　　(c)

(d)　　　　　　　　(e)　　　　　　　　(f)
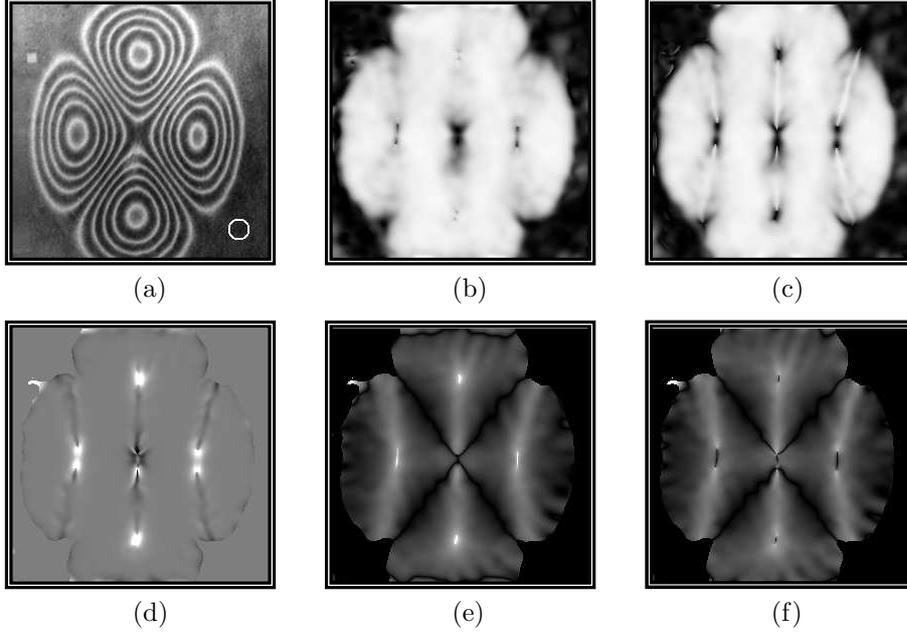
Figure 4.5: *(a) Interference pattern of vibrating plate. The superimposed circle denotes the size of the curvilinear model. (b) and (c) Confidence measures for, respectively, the circular and parabolic model (range $[0, 1]$) computed with $\sigma_g = 1.0$ and $\sigma_a = 5.0$. (d) Difference in confidence measure between circular and parabolic model (range $[-0.5, 0.5]$). (e),(f) Estimated curvatures $\hat{\kappa}$ for respectively the circular and parabolic model (log stretched).*

4.8　Appendix A

For a concentric circle model, $u(x, y, \phi, \kappa) = \sqrt{\kappa^2 w^2 + (1 - \kappa v)^2}$, the fit and residual energies are

$$
\begin{aligned}
E_f(\kappa) &= \overline{\left(\frac{(1-\kappa v)^2 f_v^2 - 2\kappa w(1-\kappa v) f_v f_w + \kappa^2 w^2 f_w^2}{(1-\kappa v)^2 + \kappa^2 w^2} \right)} \\
E_r(\kappa) &= \overline{\left(\frac{(1-\kappa v)^2 f_w^2 + 2\kappa w(1-\kappa v) f_v f_w + \kappa^2 w^2 f_v^2}{(1-\kappa v)^2 + \kappa^2 w^2} \right)}.
\end{aligned}
\tag{4.23}
$$

To obtain a closed-form solution for the curvature and the confidence measure, the local energies are computed inside a $\left(\kappa^2 w^2 + (1 - \kappa v)^2 \right)$-weighted space-variant window. This yields

$$
\begin{aligned}
\hat{E}_f &= \frac{\kappa^2 \left(\overline{v^2 f_v^2 + 2vw f_v f_w + w^2 f_w^2} \right) + 2\kappa \left(\overline{-v f_v^2 - w f_v f_w} \right) + \overline{f_v^2}}{1 - 2\kappa \overline{v} + \kappa^2 \left(\overline{v^2 + w^2} \right)} \equiv \frac{A\kappa^2 + 2B\kappa + C}{1 + D\kappa^2} \\
\hat{E}_r &= \frac{\kappa^2 \left(\overline{v^2 f_w^2 - 2vw f_v f_w + w^2 f_v^2} \right) + 2\kappa \left(\overline{-v f_w^2 + w f_v f_w} \right) + \overline{f_w^2}}{1 - 2\kappa \overline{v} + \kappa^2 \left(\overline{v^2 + w^2} \right)} \equiv \frac{E\kappa^2 + 2F\kappa + G}{1 + D\kappa^2}
\end{aligned}
\tag{4.24}
$$

(a)   (b)   (c)

(d)   (e)   (f)
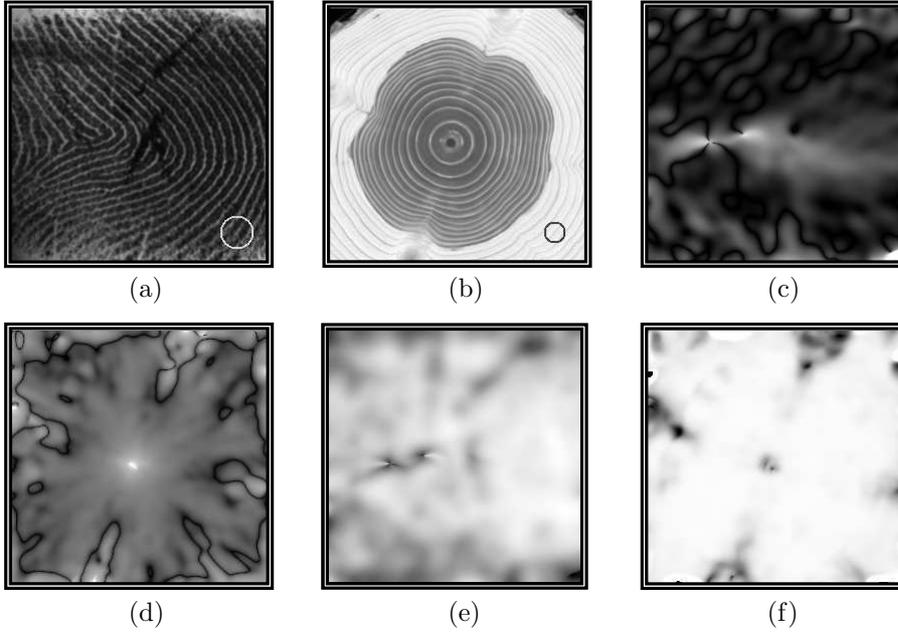
Figure 4.6: *(a) Fingerprint image. (b) CT image of trunk. (c),(d) The estimated curvature $\hat{\kappa}$ using parabolic model (log stretch) at $\sigma_g = 1.0$ and $\sigma_a = 5.0$. (e) and (f) The confidence measure of the parabolic model (range $[0, 1]$).*

with $\overline{v} = 0$. The minimization of the residual energy yields an approximation of the curvature

$$\hat{\kappa} = \frac{E - GD - \sqrt{4F^2D + (-E + GD)^2}}{2FD}. \tag{4.25}$$

The terms of $\hat{E}_f$ and $\hat{E}_r$ are expanded with Eq. (14) and

$$f_v = f_x \cos\phi + f_y \sin\phi \quad f_w = -f_x \sin\phi + f_y \cos\phi. \tag{4.26}$$

This results in

$$\begin{cases} A = \overline{x^2 f_x^2} + 2\overline{xy f_x f_y} + \overline{y^2 f_y^2} \\ B = -\left(\overline{x f_x^2} + \overline{y f_x f_y}\right)\cos\phi - \left(\overline{x f_x f_y} + \overline{y f_y^2}\right)\sin\phi \\ C = \overline{f_x^2}\cos^2\phi + 2\overline{f_x f_y}\cos\phi\sin\phi + \overline{f_y^2}\sin^2\phi \\ D = 2\sigma_a^2 \\ E = \overline{x^2 f_y^2} - 2\overline{xy f_x f_y} + \overline{y^2 f_x^2} \\ F = \left(\overline{y f_x f_y} - \overline{x f_y^2}\right)\cos\phi + \left(\overline{x f_x f_y} - \overline{y f_x^2}\right)\sin\phi \\ G = \overline{f_y^2}\cos^2\phi - 2\overline{f_x f_y}\cos\phi\sin\phi + \overline{f_x^2}\sin^2\phi \end{cases} \tag{4.27}$$

The averaged terms can be computed as global convolutions (see section on Implementation). The approximated confidence function is computed with

$$\hat{c} = \frac{\kappa^2(A - E) + 2\kappa(B - F) + (C - G)}{\kappa^2(A + E) + 2\kappa(B + F) + (C + G)}. \tag{4.28}$$

4.9   Appendix B

The terms for the parabolic confidence measure (Eq (16)) and curvature estimator (Eq. (17)) are

$$
\begin{cases}
\overline{w^2 f_w^2} = -2\left(\overline{xy f_y^2} + \overline{y^2 f_x f_y}\right)\cos^3\phi\sin\phi + \left(\overline{x^2 f_y^2} + \overline{4xy f_x f_y} + \overline{y^2 f_x^2}\right)\cos^2\phi\sin^2\phi \\
\qquad +2\left(-\overline{x^2 f_x f_y} - \overline{xy f_x^2}\right)\cos\phi\sin^3\phi + \overline{x^2 f_x^2}\sin^4\phi + \overline{y^2 f_y^2}\cos^4\phi \\
\overline{w^2 f_v^2} = -2\left(\overline{xy f_x^2} - \overline{y^2 f_x f_y}\right)\cos^3\phi\sin\phi + \left(\overline{x^2 f_x^2} - \overline{4xy f_x f_y} + \overline{y^2 f_y^2}\right)\cos^2\phi\sin^2\phi \\
\qquad +2\left(\overline{x^2 f_x f_y} - \overline{2xy f_y^2}\right)\cos\phi\sin^3\phi + \overline{x^2 f_y^2}\sin^4\phi + \overline{y^2 f_x^2}\cos^4\phi \\
\overline{w f_v f_w} = \overline{y f_x f_y}\cos^3\phi + \left(-\overline{x f_x f_y} - \overline{y\left(f_x^2 - f_y^2\right)}\right)\cos^2\phi\sin\phi \\
\qquad + \left(\overline{x\left(f_x^2 - f_y^2\right)} - \overline{y f_x f_y}\right)\cos\phi\sin^2\phi + \overline{x f_x f_y}\sin^3\phi \\
\overline{w^2} = \sigma_a^2
\end{cases}
\tag{4.29}
$$

for $\overline{f_v^2}$ and $\overline{f_w^2}$ see term C and G in appendix A.

Chapter 5

Robust Photometric Invariant Features from the Color Tensor *

5.1 Introduction

Differential-based features such as edges, corners, and salient points, are used abundantly in a variety of applications such as matching, object recognition, and object tracking [26], [64], [67]. We distinguish between feature detection and feature extraction. Feature detection aims at finding the position of features in the images, whereas for feature extraction, a position in the images is described by a set of features, which characterize the local neighborhood. Although the majority of images is recorded in color format nowadays, computer vision research is still mostly restricted restricted to luminance-based feature detection and extraction. In this chapter we focus on color information to detect and extract features.

In the basic approach to color images the gradient is computed from the derivatives of the separate channels. The derivatives of a single edge can point in opposing directions for the separate channels. DiZenzo [11] argues that a simple summation of the derivatives ignores the correlation between the channels. This also happens by converting the color image to luminance values. In the case of isoluminance of adjacent color regions it will lead to cancellation of the edge. As a solution to the opposing vector problem, DiZenzo proposes the color tensor for color gradient computation.

The same problem as occurs for color image derivatives, exists for oriented patterns (e.g. fingerprint images). Due to the high frequency nature of oriented patterns opposing derivative vectors occur in a small neighborhood. The same solution which was found for color image features, is used to compute features for oriented patterns. Kass and Witkin [37] derived orientation estimation from the structure tensor. Adap-

tations of the tensor lead to a variety of features, such as circle detectors and curvature estimation [6], [8], [24], [79]. Lee and Medioni [44] apply the structure tensor within the context of perceptual grouping.

A step forward in the understanding of color images was made by the dichromatic reflection model by Shafer [66]. The model describes how photometric changes, such as shadows and specularities, affect the RGB-values. On the basis of this model, others provided algorithms invariant to various photometric events such as shadows and specularities [17], [38]. The extension to differential photometric invariance was investigated by Geusebroek et al. [16]. In chapter 3 we introduced the photometric quasi-invariants which are a set of photometric invariant derivatives with better noise and stability characteristics compared to existing photometric invariants. Combining photometric quasi-invariants with derivative based feature detectors leads to features which can identify various physical causes, e.g. shadow corners and object corners. A drawback of the quasi-invariants is that they can only be applied for feature detection. In the case of feature extraction, where the values of multiple frames are compared, full invariance is necessary.

We propose a framework to combine the differential based-features with the photometric invariance theory. The framework is designed according to the following criteria: 1. features must target the photometric variation needed for their application. To achieve that accidental physical events, such as shadows and specularities, will not influence results. 2. features must be robust against noise and should not contain instabilities. Especially for the photometric invariant features instabilities must be dissolved. 3. physically meaningful features should be independent of the accidental choice of the color coordinate frame. Next to satisfying the criteria the framework should also be generally applicable to existing features. To meet these criteria we start from the observation that tensors are well-suited to combine first order derivatives for color images. The first contribution is a novel framework that combines tensor-based features with photometric derivatives for photometric invariant feature detection and extraction. The second contribution is that for feature extraction applications, for which quasi-invariants are unsuited, we propose a new uncertainty measure which robustifies the feature extraction. The third contribution is that the proposed features are proven to be invariant with respect to color coordinate transformations.

The chapter is organized as follows. In section 2, the prerequisites for color feature detection from tensors are discussed. In section 3, an uncertainty measure is proposed. Based on this uncertainty measure robust photometric feature extraction is derived. In section 4, a overview of tensor-based features is given. Section 5, provides several experiments and section 6 contains the concluding remarks.

5.2   Tensor-Based Features for Color Images

The extension of differential-based operations to color images can be done in various ways. The main challenge to color feature detection is how to transform the 3D-color differential structure to a representation of the presence of a feature. In this section we ensure that the transformation agrees with the criteria mentioned in the introduction.
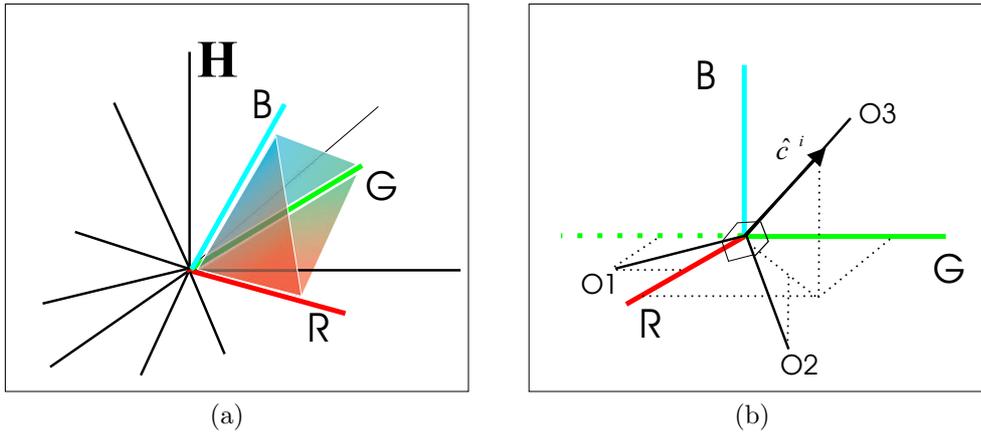
Figure 5.1: *(a) The subspace of measured light in the Hilbert space of possible spectra. (b) The RGB coordinate system and an alternative orthonormal color coordinate system which spans the same subspace (see also color plate C.9).*

In section 5.2.1 the invariance with respect to color coordinate transformation is discussed. In section 5.2.2 the transformation is written in tensor mathematics which links it with a set of tensor based features, thereby ensuring generality. In section 5.2.3 the photometric invariance of the transformation is discussed.

5.2.1 Invariance to Color Coordinate Transformations

From a physical point of view only features make sense which are invariant to rotation of the coordinate axes. This starting point has been applied in the design of image geometry features, resulting in, for example, gradient and Laplace operators [14]. For the design of physically meaningful color features not only the invariance with respect to spatial coordinate changes is desired but also the invariance with respect to color coordinate systems rotations. Features based on different measurement devices which measure the same spectral space should yield the same results.

For color images, values are represented in the *RGB* coordinate system. In fact, the infinite-dimensional Hilbert space is sampled with three probes which results in the red, green and blue channels (see Fig. 5.1). For operations on the color coordinate system to be physically meaningful they should be independent of orthonormal transformation of the three axes in Hilbert space. An example of an orthonormal color coordinate system is the opponent color space (see Fig. 5.1b). The opponent color space spans the same subspace as the subspace defined by the *RGB*-axes and hence both subspaces should yield the same features.

5.2.2 The Color Tensor

Simply summing differential structure of various color channels may result in cancellation even when evident structure exists in the image [11]. Rather than adding

the direction information (defined on $[0, 2\pi\rangle$) of the channels, it is more appropriate to sum the orientation information (defined on $[0, \pi\rangle$). Such a method is provided by tensor mathematics for which vectors in opposite directions reinforce one another. Tensors describe the local orientation rather than the direction. More precisely, the tensor of a vector and its 180° rotated counterpart vector are equal. It is for that reason that we use the tensor as a basis for color feature detection.

Given an image f, the structure tensor is given by [8]

$$\mathbf{G} = \begin{pmatrix} \overline{f_x^2} & \overline{f_x f_y} \\ \overline{f_x f_y} & \overline{f_y^2} \end{pmatrix}, \tag{5.1}$$

where the subscripts indicate spatial derivatives and the bar $(\bar{\cdot})$ indicates convolution with a Gaussian filter. Note that there are two scales involved in the computation of the structure tensor. Firstly, the scale at which the derivatives are computed and secondly the tensor-scale which is the scale at which the spatial derivatives are averaged. The structure tensor describes the local differential structure of images, and is suited to find features such as edges and corners [6], [11], [24]. For a multichannel image $\mathbf{f} = \left(f^1, f^2, ..., f^n \right)^T$, the structure tensor is given by

$$\mathbf{G} = \begin{pmatrix} \overline{\mathbf{f}_x \cdot \mathbf{f}_x} & \overline{\mathbf{f}_x \cdot \mathbf{f}_y} \\ \overline{\mathbf{f}_y \cdot \mathbf{f}_x} & \overline{\mathbf{f}_y \cdot \mathbf{f}_y} \end{pmatrix}. \tag{5.2}$$

In the case that $\mathbf{f} = (R, G, B)$, Eq. 5.2 is the color tensor. For derivatives which are accompanied with a weighting function, w_x and w_y, which appoint a weight to every measurement in $\mathbf{f_x}$ and $\mathbf{f_y}$, the structure tensor is defined by

$$\mathbf{G} = \begin{pmatrix} \dfrac{\overline{w_x^2 \mathbf{f}_x \cdot \mathbf{f}_x}}{\overline{w_x^2}} & \dfrac{\overline{w_x w_y \mathbf{f}_x \cdot \mathbf{f}_y}}{\overline{w_x w_y}} \\ \dfrac{\overline{w_y w_x \mathbf{f}_y \cdot \mathbf{f}_x}}{\overline{w_y w_x}} & \dfrac{\overline{w_y^2 \mathbf{f}_y \cdot \mathbf{f}_y}}{\overline{w_y^2}} \end{pmatrix}. \tag{5.3}$$

In section 5.2.1, we discussed that physically meaningful features should be invariant with respect to rotation of the color coordinates axes. The elements of the tensor are known to be invariant under rotation and translation of the spatial axes. To prove the invariant, we use the fact that $\frac{\partial}{\partial x} \mathbf{R} \mathbf{f} = \mathbf{R} \mathbf{f}_x$, where $\mathbf{R}$ is a rotation operator,

$$\overline{(\mathbf{R}\mathbf{f}_x)^T \mathbf{R}\mathbf{f}_y} = \overline{\mathbf{f}_x^T \mathbf{R}^T \mathbf{R}\mathbf{f}_y} = \overline{\mathbf{f}_x^T \mathbf{f}_y}. \tag{5.4}$$

where we have rewritten the inner product according to $\mathbf{f} \cdot \mathbf{f} = \mathbf{f}^T \mathbf{f}$

5.2.3 Photometric Invariant Derivatives

A good motivation for using color images is that photometric information can be exploited to understand the physical nature of features. For example, pixels can be classified as being from the same color but having different intensities which is possibly caused by a shadow or shading change in the image. Further, pixels differences can also

indicate specular reflection. For many applications it is important to distinguish the scene incidental information from material edges. When color images are converted to luminance this photometric information is lost [18].

The incorporation of photometric invariance in Eq. 5.2 can be obtained by using invariant derivatives to compute the structure tensor. In chapter 3 we derive photometric quasi-invariant derivatives and full invariant derivatives. Quasi-invariants differ from full invariants by the fact that they are variant with respect to a physical parameter. Full invariants can be computed from quasi-invariants by the normalization with a signal dependent scalar. The quasi-invariants have the advantage that they do not exhibit the instabilities common to full photometric invariants. However the applicability of the quasi-invariants is restricted to photometric invariant feature detection. For feature extraction full photometric invariance is desired.

The dichromatic model divides the reflection in the interface (specular) and body (diffuse) reflection component for optically inhomogeneous materials [66]. We assume white illumination, i.e. smooth spectrum of nearly equal energy at all wavelengths, and neutral interface reflection. For the validity of the photometric assumptions see [16], [66]. The RGB vector, $\mathbf{f} = (R, G, B)^T$, can be seen as a weighted summation of two vectors,

$$\mathbf{f} = e(m^b \mathbf{c}^b + m^i \mathbf{c}^i), \tag{5.5}$$

in which $\mathbf{c}^b$ is the color of the body reflectance, $\mathbf{c}^i$ the color of the interface reflectance (i.e. specularities or highlights), m^b and m^i are scalars representing the corresponding magnitudes of reflection and e is the intensity of the light source. For matte surfaces there is no interface reflection and the model further simplifies to

$$\mathbf{f} = em^b \mathbf{c}^b. \tag{5.6}$$

The photometric derivative structure of the image can be computed by computing the spatial derivative of Eq. 5.5

$$\mathbf{f}_x = em^b \mathbf{c}_x^b + (e_x m^b + em_x^b)\mathbf{c}^b + \left(em_x^i + e_x m^i\right) \mathbf{c}^i. \tag{5.7}$$

The spatial derivative is a summation of three weighted vectors, successively caused by body reflectance, shading-shadow and specular changes. From Eq. 5.6 it follows that for matte surfaces the shadow-shading direction is parallel to the RGB vector, $\mathbf{f}||\mathbf{c}_b$. The specular direction follows from the assumption that the color of the light source is known.

For matte surfaces (i.e. $m^i = 0$), the projection of the spatial derivative on the shadow-shading axis yields the shadow-shading variant containing all energy which could be explained by changes due to shadow and shading. Subtraction of the shadow-shading variant $\mathbf{S}_x$ from the total derivative $\mathbf{f}_x$ results in the shadow-shading quasi-invariant:

$$\begin{aligned}
\mathbf{S}_x &= \left(\mathbf{f}_x \cdot \hat{\mathbf{f}}\right) \hat{\mathbf{f}} = \left(em^b \left(\mathbf{c}_x^b \cdot \hat{\mathbf{f}}\right) + (e_x m^b + em_x^b) \left|\mathbf{c}^b\right|\right) \hat{\mathbf{f}} \\
\mathbf{S}_\mathbf{x}^\mathbf{c} &= \mathbf{f}_x - \mathbf{S}_x = em^b \left(\mathbf{c}_x^b - \left(\mathbf{c}_x^b \cdot \hat{\mathbf{f}}\right) \hat{\mathbf{f}}\right)
\end{aligned} \tag{5.8}$$

which does not contain derivative energy caused by shadows and shading. The hat, $(\hat{\ })$, denotes unit vectors. The full shadow-shading invariant results from normalizing the quasi-invariant $\mathbf{S}_x^c$ by the intensity magnitude $|\mathbf{f}|$

$$\mathbf{s}_x = \frac{\mathbf{S}_x^c}{|\mathbf{f}|} = \frac{em^b}{em^b\,|\mathbf{c}^b|}\left(\mathbf{c}_x^b - \left(\mathbf{c}_x^b\right)\cdot\hat{\mathbf{f}}\right), \qquad (5.9)$$

which is invariant for m^b.

For the construction of the shadow-shading-specular quasi-invariant, we introduce the hue-direction which is perpendicular to the light source direction $\hat{\mathbf{c}}^i$ and the shadow-shading direction $\hat{\mathbf{f}}$:

$$\hat{\mathbf{b}} = \frac{\hat{\mathbf{f}}\times\hat{\mathbf{c}}^i}{|\mathbf{f}\times\mathbf{c}^i|}. \qquad (5.10)$$

Projection of the derivative, $\mathbf{f}_x$, on the hue direction results in the shadow-shading-specular quasi-invariant:

$$\mathbf{H}_x^c = \left(\mathbf{f}_x\cdot\hat{\mathbf{b}}\right)\hat{\mathbf{b}} = em^b\left(\mathbf{c}_x^b\cdot\hat{\mathbf{b}}\right) + \left(e_x m^b + em_x^b\right)\left(\mathbf{c}^b\cdot\mathbf{b}\right). \qquad (5.11)$$

The second part of this equation is zero if we assume that shadow-shading changes do not occur within a specularity, since then either $\left(e_x m^b + em_x^b\right) = 0$ or $\left(\mathbf{c^b}\cdot\mathbf{b}\right) = \left(\mathbf{f}\cdot\mathbf{b}\right) = 0$. Subtraction of the quasi-invariant $\mathbf{H}_x^c$ from the spatial derivative $\mathbf{f}_x$ results in the shadow-shading-specular variant $\mathbf{H}_x$:

$$\mathbf{H}_x = \mathbf{f}_x - \mathbf{H}_x^c. \qquad (5.12)$$

The full shadow-shading invariant is computed by dividing the quasi-invariant by the saturation. The saturation is equal to the norm of the color-vector, $\mathbf{f}$, after the projection on the plane perpendicular to the light source direction (which is equal to subtraction of the part in the light source direction)

$$\mathbf{h}_x = \frac{\mathbf{H}_x^c}{|\mathbf{f} - (\mathbf{f}\cdot\hat{\mathbf{c}}^i)\,\hat{\mathbf{c}}^i|} = \frac{em^b}{em^b\,|\mathbf{c}^b - (\mathbf{c}^b\cdot\hat{\mathbf{c}}^i)\,\hat{\mathbf{c}}^i|}\left(\mathbf{c}_x^b\cdot\hat{\mathbf{b}}\right). \qquad (5.13)$$

The expression $\mathbf{h}_x$ is invariant for both m^i and m^b.

By projecting the local spatial derivative on three photometric axis in the RGB cube we have derived the photometric quasi-invariants. These can be combined with the structure tensor of Eq. 5.18 for photometric quasi-invariant feature detection. As discussed in section 5.2.1 we would like features to be independent of the accidental choice of the color coordinate frame. As a consequence a rotation of the color coordinates should result in the same rotation of the quasi-invariant derivatives. For example, for the shadow-shading quasi-variant $\mathbf{S}_x$ this can be proven by

$$\left(\left(\mathbf{R}\mathbf{f}_x\right)^T\mathbf{R}\hat{\mathbf{f}}\right)\left(\mathbf{R}\hat{\mathbf{f}}\right) = \left(\mathbf{f}_x^T\mathbf{R}^T\mathbf{R}\hat{\mathbf{f}}\right)\left(\mathbf{R}\hat{\mathbf{f}}\right) = \mathbf{R}\left(\mathbf{f}_x^T\hat{\mathbf{f}}\right)\hat{\mathbf{f}} = \mathbf{R}\mathbf{S}_x. \qquad (5.14)$$

Similar proofs hold for the other photometric variants and quasi-invariants. The invariance with respect to color coordinate transformation of the shadow-shading full

invariants follow from the fact that $|\mathbf{R}\mathbf{f}| = |\mathbf{f}|$. For the shadow-shading-specular full-invariant, the invariance is proven by the fact that the inner product between two vectors remains the same under rotations, and therefore $\left|\mathbf{R}\mathbf{f} - \left(\mathbf{R}\mathbf{f} \cdot \mathbf{R}\hat{\mathbf{c}}^i\right)\mathbf{R}\hat{\mathbf{c}}^i\right| = \left|\mathbf{R}\left(\mathbf{f} - \left(\mathbf{f} \cdot \hat{\mathbf{c}}^i\right)\hat{\mathbf{c}}^i\right)\right|$. Since the elements of the structure tensor are also invariant for color coordinate transformations (see Eq 5.4) the combination of the quasi-invariants and the structure tensor into a quasi-invariant structure tensor is also invariant for color coordinate transformations.

5.3 Robust Full Photometric Invariance

In section 5.2.3 the quasi- and full invariant derivatives are described. The quasi-invariants outperform the full-invariants on discriminative power and are more robust to noise (see chapter 3). However, the quasi-invariants are not suited for applications which require feature extraction. These applications compare the photometric invariant values between various images and need full photometric invariance (see Table 5.1). A disadvantage of full photometric invariants is that they are unstable in certain areas of the RGB-cube. E.g. the invariants for shadow-shading and specularities are unstable near the gray axis. These instabilities greatly reduce the applicability of the invariant derivatives for which a small deviation of the original pixel color value may result in a large deviation of the invariant derivative. In this section, we propose a measure which describes the uncertainty of the photometric invariant derivatives, thereby allowing for robust full photometric invariant feature detection.

We will first derive the uncertainty for the shadow-shading full invariant from its relation to the quasi-invariant. We assume additive uncorrelated uniform Gaussian noise. Due to the high-pass nature of differentiation we assume the noise of the zero order signal ($|\mathbf{f}|$) to be negligible compared to the noise on the first order signal ($\mathbf{S}_x^c$). In section 5.2.3, the quasi-invariant has been derived by a linear projection of the derivative $\mathbf{f}_x$ on the plane perpendicular to the shadow-shading direction. Therefore, uniform noise in $\mathbf{f}_x$ will result in uniform noise in $\mathbf{S}_x^c$. The noise in the full invariant can be written as

$$\tilde{\mathbf{s}}_x = \frac{\mathbf{S}_x^c + \sigma}{|\mathbf{f}|} = \frac{\mathbf{S}_x^c}{|\mathbf{f}|} + \frac{\sigma}{|\mathbf{f}|}. \tag{5.15}$$

The uncertainty of the measurement of $\tilde{\mathbf{s}}_x$ depends on the magnitude of $|\mathbf{f}|$. For small $|\mathbf{f}|$ the error increases proportionally. Therefore we propose to weight the full shadow-shading invariant with the function $w = |\mathbf{f}|$ to robustify the color tensor-based on the chromatic invariant. For shadow-shading invariance examples of the equations used to compute the color tensor are given in Table 5.1.

For the shadow-shading-specular invariant, the weighting function should be proportional with the saturation, since

$$\tilde{\mathbf{h}}_x = \frac{\mathbf{H}_x^c + \sigma}{|\mathbf{s}|} = \frac{\mathbf{H}_x^c}{|\mathbf{s}|} + \frac{\sigma}{|\mathbf{s}|}. \tag{5.16}$$

This leads us to propose $w = |\mathbf{s}|$ as the weighting function of the hue derivative $\tilde{\mathbf{h}}_x$ see Fig. 5.2). On place where there is an edge, the saturation drops, and with the

	detection	extraction
quasi invariant	+++	-
full invariant	+	+
robust full invariant	++	++

Table 5.1: *Applicability of the different invariants for feature detection and extraction.*

saturation the certainty of the hue measurement. The quasi-invariant (see Fig. 5.2d), which is equal to the weighted hue, is more stable than the full invariant derivative due to the incorporation of the certainty in the measurements. With the derived weighting function we can compute the robust photometric invariant tensor (Eq. 5.3).

The uncertainties of the full-invariant by ways of error-propagation have also been investigated by Stokman and Gevers[19]. Our assumption of uniform noise in the RGB channels together with the choice of invariants based on orthogonal color space transformations leads to a simplification of the uncertainty measure. It also connects with the intuitive notion that the uncertainty of the hue is depended on the saturation and the uncertainty of the chromaticity (shadow-shading invariant) with the intensity.

5.4   Color Tensor-Based Features

In this section we show the generality of the proposed method by summing features which can be derived from the color tensor. In section 5.2.3 and in section 5.3 we described how to compute invariant derivatives. Dependent on the task at hand we proposed to use either quasi-invariants for detection or robust full invariants for extraction. The features in this chapter will be derived for $\mathbf{g}_x$. By replacing the inner
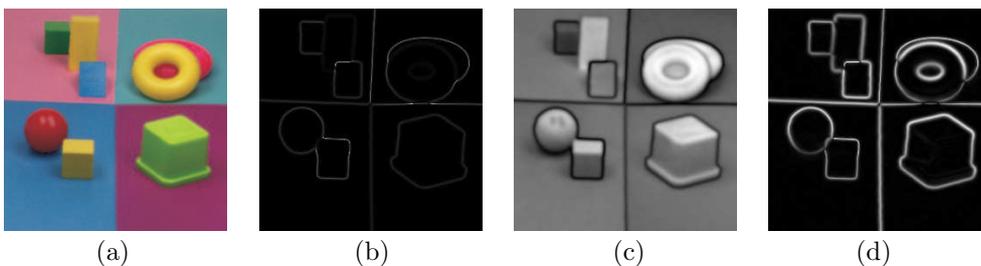

Figure 5.2: *(a) test image (b) hue derivative (c) saturation (d) quasi-invariant (see also color plate C.10).*

product of $\mathbf{g}_x$ by one of the following

$$\left\{ \overline{\mathbf{f}_x \cdot \mathbf{f}_x}, \overline{\mathbf{S}_x^c \cdot \mathbf{S}_x^c}, \frac{\overline{\mathbf{S}_x^c \cdot \mathbf{S}_x^c}}{|\mathbf{f}|^2}, \overline{\mathbf{H}_x^c \cdot \mathbf{H}_x^c}, \frac{\overline{\mathbf{H}_x^c \cdot \mathbf{H}_x^c}}{|\mathbf{s}|^2} \right\}. \tag{5.17}$$

the acquired photometric invariant features are attained. In section 5.4.1 we describe features derived from the eigenvalues of the tensor. In section 5.4.2 features which are derived from an adapted version of the structure tensor and in section 5.4.3 we describe color optical flow.

5.4.1   Eigenvalue-Based Features

Eigenvalue analysis of the tensor leads to two eigenvalues which are defined by

$$\begin{aligned} \lambda_1 &= \tfrac{1}{2} \left(\overline{\mathbf{g}_x \cdot \mathbf{g}_x} + \overline{\mathbf{g}_y \cdot \mathbf{g}_y} + \sqrt{(\overline{\mathbf{g}_x \cdot \mathbf{g}_x} - \overline{\mathbf{g}_y \cdot \mathbf{g}_y})^2 + (2\overline{\mathbf{g}_x \cdot \mathbf{g}_y})^2} \right) \\ \lambda_2 &= \tfrac{1}{2} \left(\overline{\mathbf{g}_x \cdot \mathbf{g}_x} + \overline{\mathbf{g}_y \cdot \mathbf{g}_y} - \sqrt{(\overline{\mathbf{g}_x \cdot \mathbf{g}_x} - \overline{\mathbf{g}_y \cdot \mathbf{g}_y})^2 + (2\overline{\mathbf{g}_x \cdot \mathbf{g}_y})^2} \right). \end{aligned} \tag{5.18}$$

The direction of λ_1 indicates the prominent local orientation

$$\theta = \tfrac{1}{2} \arctan \left(\frac{2\overline{\mathbf{g}_x \cdot \mathbf{g}_y}}{\overline{\mathbf{g}_x \cdot \mathbf{g}_x} - \overline{\mathbf{g}_y \cdot \mathbf{g}_y}} \right). \tag{5.19}$$

The λ's can be combined to give the following local descriptors:

- $\lambda_1 + \lambda_2$ describes the total local derivative energy.

- λ_1 is the derivative energy in the most prominent direction.

- $\lambda_1 - \lambda_2$ describes the line-energy (see [62]). The derivative energy in the prominent orientation is corrected for the energy contributed by the noise λ_2.

- λ_2 describes the amount of derivative energy perpendicular to the prominent local orientation which is used to select features for tracking [67].

An often applied feature detector is the Harris corner detector [27]. The color Harris operator H can be written as a function of the eigenvalues of the structure tensor

$$\begin{aligned} H\mathbf{f} &= \overline{\mathbf{g}_x \cdot \mathbf{g}_x} \, \overline{\mathbf{g}_y \cdot \mathbf{g}_y} - \overline{\mathbf{g}_x \cdot \mathbf{g}_y}^2 - k \left(\overline{\mathbf{g}_x \cdot \mathbf{g}_x} + \overline{\mathbf{g}_y \cdot \mathbf{g}_y} \right)^2 \\ &= \lambda_1 \lambda_2 - k \left(\lambda_1 + \lambda_2 \right)^2. \end{aligned} \tag{5.20}$$

5.4.2   Adaptations of the Color Tensor

The same equations as DiZenzo's equations for orientation estimation are found by Kass and Witkin [37]. They studied orientation estimation for oriented patterns (e.g. fingerprint images). Oriented patterns are defined as patterns with a dominant orientation everywhere. For oriented patterns other mathematics are needed than

for regular object images. The local structure of object images is described by a step edge, whereas for oriented patterns the local structure is described as a set of lines (roof edges). Lines generate opposing vectors on a small scale. Hence for geometric operations on oriented patterns, methods are needed for which opposing vectors enforce one another. This is the same problem as encountered for all color images, where the opposing vector problem does not only occur for oriented patterns, but also for step edges, for which the opposing vectors occur in the different channels. Hence similar equations were found in both fields. Next to orientation estimation, a number of other estimators were proposed by oriented pattern research [6], [24], [79]. These operation are based on adaptations of the structure tensor and can also be applied to the color tensor.

The structure tensor of Eq. 5.2 can also be seen as a local projection of the derivative energy on two perpendicular axes, namely $\mathbf{u}_1 = \begin{pmatrix} 1 & 0 \end{pmatrix}^T$ and $\mathbf{u}_2 = \begin{pmatrix} 0 & 1 \end{pmatrix}^T$,

$$\mathbf{G}^{\mathbf{u}_1,\mathbf{u}_2} = \begin{pmatrix} \overline{(\mathbf{G}_{x,y}\mathbf{u}_1) \cdot (\mathbf{G}_{x,y}\mathbf{u}_1)} & \overline{(\mathbf{G}_{x,y}\mathbf{u}_1) \cdot (\mathbf{G}_{x,y}\mathbf{u}_2)} \\ \overline{(\mathbf{G}_{x,y}\mathbf{u}_1) \cdot (\mathbf{G}_{x,y}\mathbf{u}_2)} & \overline{(\mathbf{G}_{x,y}\mathbf{u}_2) \cdot (\mathbf{G}_{x,y}\mathbf{u}_2)} \end{pmatrix} \tag{5.21}$$

in which $\mathbf{G}_{x,y} = \begin{pmatrix} \mathbf{g}_x & \mathbf{g}_y \end{pmatrix}$. From the Lie group of transformation several other choices of perpendicular projections can be derived [6], [24]. They include feature extraction for circle, spiral and star-like structures.

The star and circle detector is given as an example. It is based on $\mathbf{u}_1 = \frac{1}{\sqrt{x^2+y^2}} \begin{pmatrix} x & y \end{pmatrix}^T$ which coincide with the derivative pattern of a circular patterns and $\mathbf{u}_2 = \frac{1}{\sqrt{x^2+y^2}} \begin{pmatrix} -y & x \end{pmatrix}^T$ which denotes the perpendicular vector field which coincides with the derivative pattern of starlike patterns. These vectors can be used to compute the adapted structure tensor with Eq. 5.21. Only the elements on the diagonal have non zero entries and are equal to

$$\mathbf{H} = \begin{pmatrix} \overline{\frac{x^2}{x^2+y^2}\mathbf{g}_x\cdot\mathbf{g}_x} + \overline{\frac{2xy}{x^2+y^2}\mathbf{g}_x\cdot\mathbf{g}_y} + \overline{\frac{y^2}{x^2+y^2}\mathbf{g}_y\cdot\mathbf{g}_y} & 0 \\ 0 & \overline{\frac{x^2}{x^2+y^2}\mathbf{g}_y\cdot\mathbf{g}_y} - \overline{\frac{2xy}{x^2+y^2}\mathbf{g}_x\cdot\mathbf{g}_y} + \overline{\frac{y^2}{x^2+y^2}\mathbf{g}_x\cdot\mathbf{g}_x} \end{pmatrix}, \tag{5.22}$$

here λ_1 describes the amount of derivative energy contributing to circular structures and λ_2 the derivative energy which describes a starlike structure. Similar to the proof given in Eq. 5.4 the elements of Eq. 5.22 can be proven to be invariant under transformations of the RGB-space.

Curvature is another feature which can be derived from an adaption of the structure tensor, as discussed in chapter 4. The fit between the local differential structure and a parabolic model function can be written as a function of the curvature. Finding the optimum of this function yields an estimation of the local curvature. For vector data the equation for the curvature is given by

$$\kappa = \frac{\overline{w^2\mathbf{g}_v\cdot\mathbf{g}_v} - \overline{w^2}\cdot\overline{\mathbf{g}_w\cdot\mathbf{g}_w} - \sqrt{\left(\overline{w^2}\cdot\overline{\mathbf{g}_w\cdot\mathbf{g}_w} - \overline{w^2\mathbf{g}_v\cdot\mathbf{g}_v}\right)^2 + 4\overline{w^2}\cdot\overline{w\mathbf{g}_v\cdot\mathbf{g}_w}^2}}{2\overline{w^2}\cdot\overline{w\mathbf{g}_v\cdot\mathbf{g}_w}} \tag{5.23}$$

in which $\mathbf{g}_v$ and $\mathbf{g}-_w$ are the derivatives in gauge coordinates.

5.4.3 Color Optical Flow

Optical flow can also be computed from the structure tensor. This is originally proposed by Simoncelli [68] and has been extended to color in [4], [21]. The vector of a multi-channel point over time stays constant [31], [48]

$$\frac{d\mathbf{g}}{dt} = \mathbf{0}. \tag{5.24}$$

Differentiating yields the following set of equations

$$\mathbf{G_{x,y}}\,\mathbf{v} + \mathbf{g_t} = \mathbf{0} \tag{5.25}$$

with $\mathbf{v}$ the optical flow. To solve the singularity problem and to robustify the optical flow computation we follow Simoncelli [68] and assume a constant flow within a Gaussian window. Solving Eq. 5.25 leads to the following optical flow equation

$$\mathbf{v} = \overline{(\mathbf{G}_{x,y} \cdot \mathbf{G}_{x,y})^{-1}}\;\overline{\mathbf{G}_{x,y} \cdot \mathbf{g}}_t = \mathbf{M}^{-1}\mathbf{b} \tag{5.26}$$

with

$$\mathbf{M} = \left(\begin{array}{cc} \overline{\mathbf{g}_x \cdot \mathbf{g}_x} & \overline{\mathbf{g}_x \cdot \mathbf{g}_y} \\ \overline{\mathbf{g}_y \cdot \mathbf{g}_x} & \overline{\mathbf{g}_y \cdot \mathbf{g}_y} \end{array} \right) \tag{5.27}$$

and

$$\mathbf{b} = \left(\begin{array}{c} \overline{\mathbf{g}_x \cdot \mathbf{g}_t} \\ \overline{\mathbf{g}_y \cdot \mathbf{g}_t} \end{array} \right). \tag{5.28}$$

The assumption of color optical flow based on RGB is that RGB pixel values remain constant over time (see Eq. 5.24). A change of brightness introduced due to a shadow, or a light source with fluctuating brightness such as the sun results in non existent optical flow. This problem can be overcome by assuming constant chromaticity over time. For photometric invariant optical flow, full invariance is necessary since the optical flow estimation is based upon comparing the (extracted) edge response of multiple frames. Consequently photometric invariant optical flow can be attained by replacing the inner product of $\mathbf{g}_x$ by one of the following

$$\left\{ \frac{\overline{\mathbf{S}_x^c \cdot \mathbf{S}_x^c}}{|\mathbf{f}|^2}, \frac{\overline{\mathbf{H}_x^c \cdot \mathbf{H}_x^c}}{|\mathbf{s}|^2} \right\}. \tag{5.29}$$

5.5 Experiments

The experiments test the features on the required criteria of our framework: 1. photometric invariance, 2. robustness. The third criterium, i.e. invariance with respect to color coordinate transformations, we have already proven theoretically. In this section we aim to demonstrate invariance by experiment and illustrate the generality of the experiments by the variety of examples. For all experiments the derivatives are

	detection error %		extraction error %	
std. noise	5	20	5	20
shadow-shading				
quasi invariant	<u>5.1</u>	<u>20.2</u>	100	100
full invariant	11.7	50.1	8.7	56.6
robust full invariant	6.4	37.7	<u>3.0</u>	<u>35.3</u>
shadow-shading-specular				
quasi invariant	<u>9.7</u>	<u>46.6</u>	100	98.2
full invariant	38.8	75.5	62.3	84.0
robust full invariant	15.7	60.2	<u>9.8</u>	<u>66.6</u>

Table 5.2: *Percentage of falsely detected points and percentage of wrongly classified points. Classification is based on the extraction of invariant information. Uncorrelated Gaussian noise is added with standard deviation 5 and 20.*

computed with a Gaussian derivative of $\sigma = 1$ and the color tensor scale is computed with $\sigma = 3$, except when mentioned otherwise. The experiments are performed using a Sony 3CCD color camera XC-003P, Matrox Corona Frame-grabber, and two Osram 18 Watt "Lumilux deLuxe daylight" fluorescent light sources.

5.5.1   Photometric Invariant Harris Point Detection

Robustness with respect to photometric changes, stability of the invariants, and robustness to noise, are tested. Further the ability of invariants to detect and extract


| (a) | (b) | (c) |
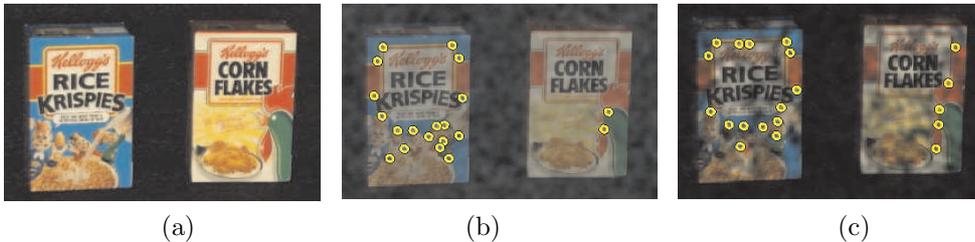
Figure 5.3: *(a) An example from Soil-47 image. (b) shadow-shading distortion with the shadow-shading quasi-invariant Harris points superimposed (c) specular distortion and the shadow-shading-specular Harris points superimposed (see also color plate C.11).*

features is examined, see also Table 5.1. The experiment is performed with the photo-metric invariant Harris corner detector (Eq. 5.20) and is executed on the Soil47 multi object set [42], which comprises of 23 images, see Fig. 5.3a.

First, the feature detection accuracy of the invariants is tested. For each image and invariant, the 20 most prominent Harris points are extracted. Next, Gaussian uncorrelated noise is added to the data, and the Harris point detection is computed 10 times per image. The percentage of points which do not correspond to the Harris points in the noiseless case are given in Table 5.2. The Harris point detector based on the quasi-invariant outperforms the alternatives. The instability within the full invariant can be partially repaired by the robust full invariant, however for detection purposes the quasi-invariants remain the best choice.

Next, the feature extraction for the invariants is tested. Again the 20 most promi-nent Harris points are detected in the noise free image. For these points the pho-tometric invariant derivative energy is extracted by $\sqrt{\lambda_1 + \lambda_2 - 2\lambda_n}$, where λ_n is an estimation of the noise which contributes to the energy in both λ_1 and λ_2. To imitate photometric variations of images we apply the following photometric distortion to the images (compare with Eq. 5.5)

$$\mathbf{g}(\mathbf{x}) = \alpha(\mathbf{x})\mathbf{f}(\mathbf{x}) + \beta(\mathbf{x})\mathbf{c}^i + \eta(\mathbf{x}),\qquad(5.30)$$

where $\alpha(\mathbf{x})$ is a smooth function resembling variation similar to shading and shadow effects, $\beta(\mathbf{x})$ is a smooth function which imitates specular reflections, and $\eta(\mathbf{x})$ is Gaussian noise. To test the shadow-shading extraction $\alpha(\mathbf{x})$ is chosen to vary be-tween 0 and 1, and $\beta(\mathbf{x})$ is 0. To test the shadow-shading-specular invariants $\alpha(\mathbf{x})$ was chosen constant at 0.7 and $\beta(\mathbf{x})$ varied between zero and fifty. After the photo-metric distortion the derivative energy is extracted at the same twenty points. The extraction is considered correct if the deviation of the derivative energy between the distorted and the noise-free case is less then 10 percent. The results are given in Table 5.2. Quasi-invariants which not suited for extraction have a hundred percent error. The full invariants have better results but with worsening signal-to-noise ratio its performance drops drastically. In accordance with the theory in section 5.3 the robust full invariants successfully improve the performance.

5.5.2 Color Optical Flow

Robustness of the full photometric invariance features is tested on photometric in-variant optical flow estimation. The optical flow is estimated on a synthetical image sequence with constant optical flow. We use the robust full photometric structure tensor for the estimation of optical flow and compare it with 'classical' photometric optical flow as proposed by [21]. Derivatives are computed with a Gaussian derivative of $\sigma = 1$ and the color tensor scale is $\sigma = 5$.

The shadow-shading photometric optical flow is tested on image with decreasing intensity (see Fig. 5.4a) which is shifted one pixel per frame. Uncorrelated Gaussian noise with $\sigma = 20$ is added to the sequence. In Fig. 5.4b,c the mean and the standard deviation of the optical flow along the y-axis of Fig. 5.4a are depicted. Similarly to the shadow-shading-specular invariant optical flow is tested on a sequence with
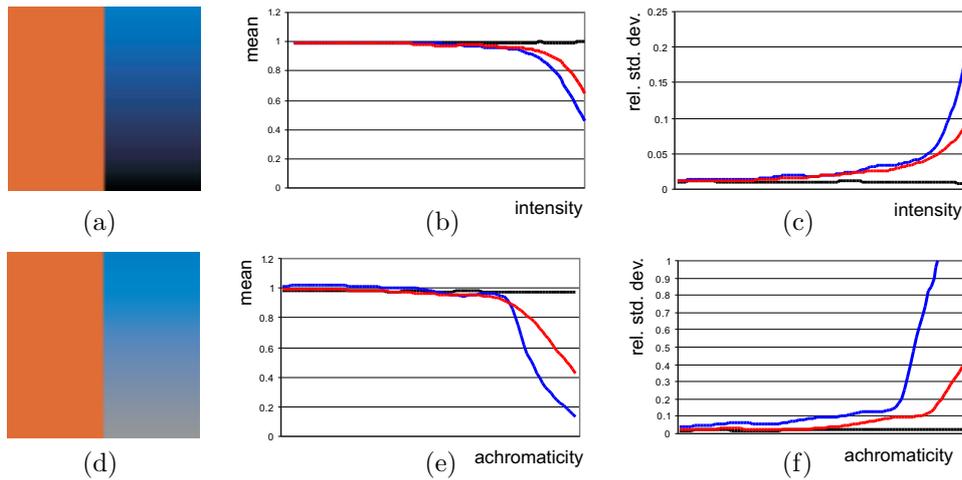
Figure 5.4: *(a),(d) frame from test sequence with constant optical flow of one pixel per frame. (b),(c) mean and relative standard deviation mean of the optical flow based on RGB (black line), shadow-shading invariant (blue line) and robust shadow-shading invariant (red line). (e),(f) mean and relative standard deviation of the optical flow based on RGB (black line), shadow-shading-specular invariant (blue line) and robust shadow-shading-specular invariant (red line) (see also color plate C.12).*

increasing achromaticity along the axes (see Fig. 5.4d,e,f.). The results show that robust invariant methods (red lines) outperform the standard photometric optical flow (blue lines). The gained robustness becomes apparent for the measurements around the instable region. Which are the black area for the shadow-shading invariant and the achromatic, grey area for the shadow-shading-specular invariant optical flow.

As an example of a real-world scene, multiple frames are taken from static objects while the light source position is changed. This results in a violation of the brightness
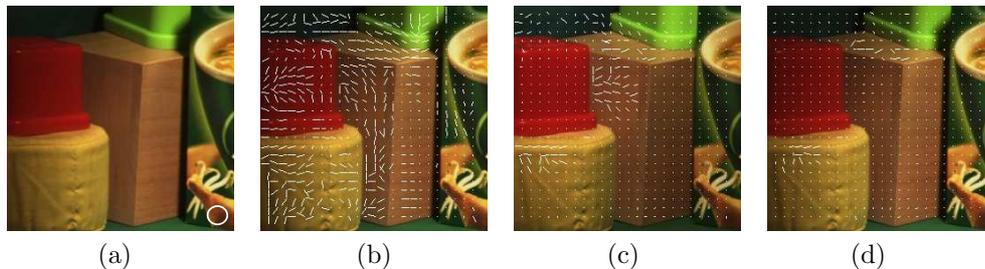

Figure 5.5: *(a) frame 1 of object scene with filter size superimposed on it. (b) RGB gradient optical flow (c) shadow-shading invariant optical flow and (d) robust shadow-shading invariant optical flow (see also color plate C.13).*
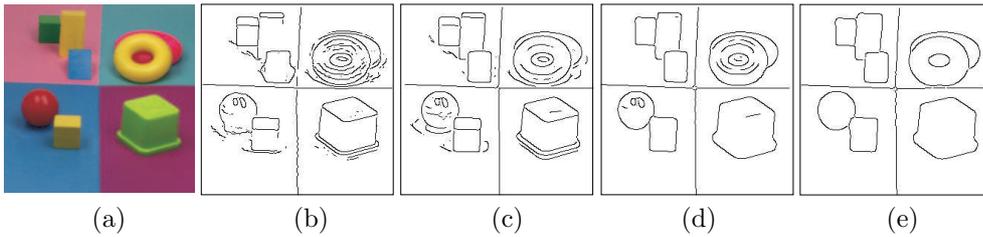
Figure 5.6: *(a) input image with Canny edge detection based on successively (b) luminance derivative (c) RGB derivatives (d) the shadow-shading quasi-invariant (e) the shadow-shading-specular quasi-invariant (see also color plate C.14).*

constraint by changing shading and moving shadows. Since both the camera and the objects did not move the ground truth optical flow is zero. The violation of the brightness constraint disturbs the optical flow estimation based on the RGB (Fig. 5.5b). The shadow-shading invariant optical flow estimation is much less disturbed by the violation of the brightness constrain (Fig. 5.5c). However, the flow estimation is still unstable around some of the edges. The robust shadow-shading invariant optical flow has the best results and is only unstable in low-gradient area's (Fig. 5.5d).

5.5.3 Color Canny Edge Detection

We illustrate the use of eigenvalue-based features by adapting the Canny edge detection algorithm to allow for vectorial input data. The algorithm consists of the following steps

1. Compute the spatial derivatives, $\mathbf{f}_x$, and combine them if desired into a quasi-invariant (Eq. 5.8 or Eq. 5.11).

2. Compute the maximum eigenvalue (Eq. 5.18) and its orientation (Eq. 5.19).

3. Apply non-maximum suppression on λ_1 in the prominent direction.

In Fig. 5.6 the results of color Canny edge detection for several photometric quasi-invariants is shown. The results show that the luminance-based Canny, Fig. 5.6b, misses several edges which are correctly found by the RGB-based method , Fig. 5.6c. Also the removal of spurious edges by photometric invariance is demonstrated. In Fig. 5.6d the edge detection is robust to shadow and shading changes and only detects material and specular edges. In Fig. 5.6e only the material edges are depicted.

5.5.4 Circular Object Detection

The use of photometric invariant orientation and curvature estimation is demonstrated on a circle detection example. Other than the previous experiments these images have been recorded by the Nikon Coolpix 950, a commercial digital camera of

(a) (b) (c)
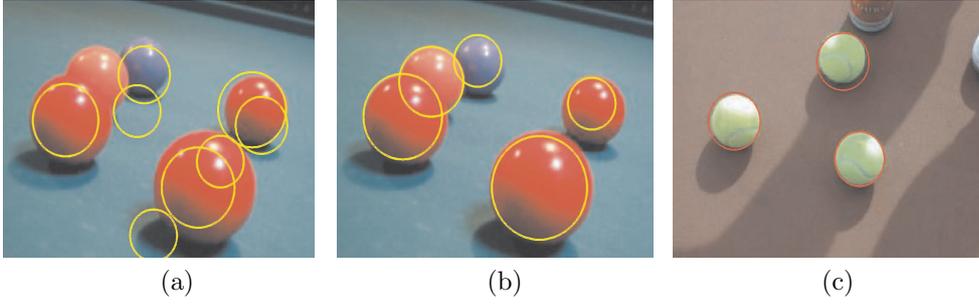
Figure 5.7: *(a) detected circles based on luminance (b) detected circles based on shadow-shading-specular quasi-invariant (c) detected circles based on shadow-shading-specular quasi-invariant (see also color plate C.15).*

average quality. The images have size 267x200 pixels with JPEG compression. The digitization was done in 8 bits per color.

Circular object recognition is complicated due to shadow, shading and specular events which influence the feature extraction. We apply the following algorithm for circle detection

1. Compute the spatial derivatives, $\mathbf{f}_x$, and combine them if desired into a quasi-invariant (Eq. 5.8 or Eq. 5.11).

2. Compute the local orientation, Eq. 5.19, and curvature, Eq. 5.23.

3. Compute the hough space [3], $H\left(R, x^0, y^0\right)$, where R is the radius of the circle and x^0 and y^0 indicate the center of the circle. The computation of the orientation and curvature reduces the number of votes per pixel to one. Namely, for a pixel at position $\mathbf{x} = \left(x^1, y^1\right)$,

$$
\begin{aligned}
R &= \tfrac{1}{\kappa} \\
x^0 &= x^1 + \tfrac{1}{\kappa}\cos\theta \\
y^0 &= y^1 + \tfrac{1}{\kappa}\sin\theta.
\end{aligned}
\tag{5.31}
$$

Every pixel votes with its the derivative energy $\sqrt{\mathbf{f}_x \cdot \mathbf{f}_x}$.

4. Compute the maxima in the hough space. These maxima indicate the circle centers and the radii of the circle.

In Fig. 5.7 the results of the circle detection are given. The luminance-based circle detection is corrupted by the photometric variation in the image. Nine circles had to be detected before the five balls were detected. For the shadow-shading-specular quasi-invariant based method the five most prominent peaks in the hough space coincide with reasonable estimates of the radii and center points of the circles. Note that, although the recordings do not fulfill the assumptions on which the dichromatic model is based, such as white light source, saturated pixels and linear camera response, the
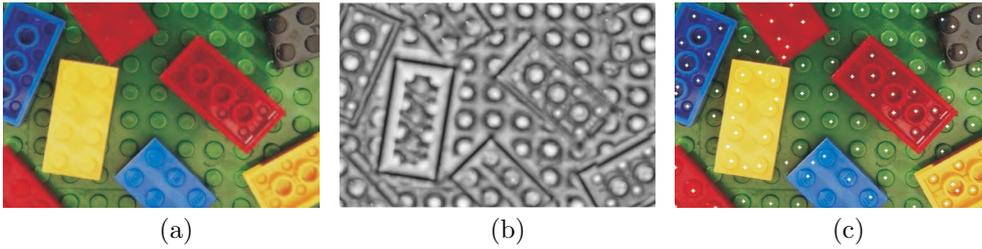
Figure 5.8: *(a) input image (b) the circularity coefficient C (c) the detected circles (see also color plate C.16).*

invariants still improve performance by partially suppressing scene incidental events, such as shadows and specularities. In Fig. 5.7 an outdoor example with a shadow partially covering the objects is given.

5.5.5 Local Color Symmetry Detector

The applicability of the features derived from an adaptation of the structure tensor (section 5.4.2) is illustrated here for a symmetry detector. We apply the circle detector to an image containing Lego-blocks (Fig. 5.8). Because we know that the color within the blocks remains the same, the circle detection is done on the shadow-shading-specular variant, $\mathbf{H}_x$ (Eq. 5.11). The shadow-shading-specular variant contains all the derivative energy except for the energy which can only be caused by a material edge. With the shadow-shading-specular variant the circular energy λ_1 and the starlike energy λ_2 are computed according to Eq. 5.22. Dividing the circular energy by the total energy yields a descriptor of local circularity (see Fig. 5.8b)

$$C = \frac{\lambda_1}{\lambda_1 + \lambda_2}. \tag{5.32}$$

The superimposed maxima of C, Fig. 5.8c, give good estimation of the circle centers.

5.6 Conclusions

In this chapter we proposed a framework to combine tensor-based features and photometric invariance theory. The tensor basis of these features ensures that opposing vectors in different channels do not cancel out, but instead that they reinforce each other. To overcome the instability caused by transformation to an photometric full invariant, we propose an uncertainty measure to accompany the full invariant. This uncertainty measure is incorporated in the color tensor to generate robust photometric invariant features. Experiments show that: 1) the color based features outperform their luminance counterparts, 2) the quasi-invariants give stable detection, and 3) that the robust invariants give better extraction results.

Chapter 6

Boosting Color Saliency in Image Feature Detection

6.1 Introduction

Indexing objects and object categories as an ordered collection of salient points has been successfully applied to image matching, content-based retrieval, learning and recognition [13], [47], [54], [63], [65], [85]. Salient points are local features in the image which exhibit geometrical structure, such as T-junctions, corners, and symmetry points. The aim of salient point detection is to represent objects more concisely and robust to varying viewing conditions, such as changes due to zooming, rotation, and illumination changes. Applications based on salient points are generally composed of three phases: 1. a feature detection phase locating the features. 2. an extraction phase in which local descriptions are extracted at the detected locations and 3. a matching phase in which the extracted descriptors are matched against a database of descriptors. In this chapter, the focus is to improve the salient point detection phase.

Although the majority of image data is in color format nowadays, most salient point detectors are still luminance based. They typically focus on shape saliency rather than color saliency [44], [93]. They focus on corner points without distinguishing low-salient black-and-white corners from high-salient red-green corners. Only recently color information has been incorporated in the detection phase. Montesinos et al. [55] propose an extension of the luminance Harris corner detector to color [27]. Heidemann [29] incorporates color into the generalized symmetry transform proposed by Reisfeld et al. [60]. Both methods achieve a performance gain for near isoluminant events. However, since the luminance axis remains the major axes of color variation in the RGB-cube, results do not differ greatly from luminance based feature detection. Itti et al. [33] use color contrast as a clue for salience. Their method is based on a zero-order signal which is not easily extendable to differential-based features.

For the evaluation of salient point detectors Schmid et al. [64] proposed two criteria: 1. *repeatability*, salient point detection should be stable under the varying viewing

conditions, such as geometrical changes and photometric changes. 2. *distinctiveness*, salient points should focus on events with a low probability of occurrence. Most salient point detectors are designed according to these criteria. They focus on two dimensional structures, such as corners, which are stable and distinctive at the same time. Although color is believed to play an important role in attributing saliency [34], the explicit incorporation of color distinctiveness into the design of salient points detectors has, to our knowledge, not been done.

A remarkable phenomenon appears when studying the statistics of color image derivatives. In histograms of color derivatives, points of equal frequency form regular structures. These color image derivatives play two roles in salient point detection. Firstly, they are input to the saliency function, which based on local derivatives probes for salient structures. Secondly, they are part of the extracted local features, on which the distinctiveness of the salient point detector is based. This double role, together with the statistical finding described above, leads to the following question: How can we exploit the regularity of the distinctiveness of color image derivatives to improve salient feature detection ?

In this chapter we aim to incorporate color distinctiveness into salient point detection. The extension should be general and hence be easy to incorporate in existing salient point detectors. For a color image, with values $\mathbf{f} = (R, G, B)^T$, salient points are the maxima of the saliency map, which compares the derivative vectors in a neighborhood fixed by scale σ,

$$s = H^\sigma \left(\mathbf{f}_x, \mathbf{f}_y \right) \tag{6.1}$$

where H is the saliency function and the subscript indicates differentiation with respect to the parameter. This type of saliency maps include [6], [27], [29], [44], [76]. The impact of a derivative vector on the outcome of the local salience depends on its vector norm, $|\mathbf{f}_x|$. Hence, vectors with equal norm have an equal impact on the local saliency. Rather than deriving saliency from the vector norm, the challenge is to adapt the saliency function in order that vectors with equal color distinctiveness have equal impact on the saliency function.

6.2   Color Distinctiveness

The efficiency of salient point detection depends on the distinctiveness of the extracted salient points. At the salient points' positions, local neighborhoods are extracted and described by local image descriptors. The distinctiveness of the descriptor defines the conciseness of the representation and the discriminative power of the salient points. The distinctiveness of interest points is measured by its information content [64].

For luminance-based descriptors, the information content is measured by looking at the distinctiveness of the differential invariants described by the local 2-jet [40] at the detected points [63]. Montesinos et al. [55] argue that due to the extra information available in color images the color 1-jet is sufficient for local structure description. The color 1-jet descriptor is given by

$$\mathbf{v} = \begin{pmatrix} R & G & B & R_x & G_x & B_x & R_y & G_y & B_y \end{pmatrix}^T. \tag{6.2}$$

The information content of this color descriptor includes the information content of more complex local color descriptors such as color differential invariant descriptors, since these complex descriptors are computed from the elements of Eq. 6.2.

From information theory it is known that the information content of an event is dependent on its frequency or probability. Events which occur rarely are more informative. The dependency of information content on its probability is given by

$$I\left(\mathbf{v}\right) = -\log\left(p\left(\mathbf{v}\right)\right) \tag{6.3}$$

where $p\left(\mathbf{v}\right)$ is the probability of the descriptor $\mathbf{v}$. The information content of the descriptor, given by Eq. 6.2, is approximated by assuming independent probabilities of the zeroth order signal and the first order derivatives

$$p\left(\mathbf{v}\right) = p\left(\mathbf{f}\right)p\left(\mathbf{f}_x\right)p\left(\mathbf{f}_y\right). \tag{6.4}$$

To improve the information content of the salient point detector, defined by Eq. 6.1, the probability of the derivatives, $p\left(\mathbf{f}_x\right)$, should be small.

We can now restate the aim of this chapter in a more precise manner. The aim is to find a transformation $g : \Re^3 \rightarrow \Re^3$ for which holds that

$$p\left(\mathbf{f}_x\right) = p\left(\mathbf{f}_x^{'}\right) \leftrightarrow |g\left(\mathbf{f}_x\right)| = \left|g\left(\mathbf{f}_x^{'}\right)\right|. \tag{6.5}$$

This implies that vectors with equal information content have equal impact on the saliency function. The transformation, attained by the function g, is called *color saliency boosting*. Similar equations hold for $p\left(\mathbf{f}_y\right)$. Once a color boosting function g has been found, the color boosted saliency can be computed with

$$s = H^{\sigma}\left(g\left(\mathbf{f}_x\right), g\left(\mathbf{f}_y\right)\right). \tag{6.6}$$

The saliency map which used to derive saliency from the orientations and gradient strength of the derivatives in a local neighborhood, is after color boosting based on the orientations and the information content of these derivatives. Gradient strength has been replaced by information content, thereby aiming for higher information content.

From Eq. 6.5 the color boosting function g is found by analyzing the probabilities of the derivatives. The channels of $\mathbf{f}_x$, $\{R_x, G_x, B_x\}$ are correlated due to the physics of the world. Photometric events in the world, such as shading, and reflection of the light source in specularities influence RGB values in a well defined manner [66]. Before investigating the statistics of color derivatives, the derivatives need to be transformed to a color space which is uncorrelated with respect to these photometric events.

6.3 Physics-Based Decorrelation

Here we describe three color coordinate transformations which partition RGB-space differently. The transformation are derived from photometric invariance theory [66]. Photometric invariance theory allows us to distinguish between various photometric

causes for features in the image, such as shadows, shading, specularities and object re-
flectance changes. The theory is based on the dichromatic reflection model introduced
by Shafer [66]. Geusebroek et al. [16] extended the photometric reflection theory to
differential-based operations. In chapter 3 we introduced the quasi-invariant deriva-
tives to improve noise characteristics. Here we use the same color transformations to
decorrelate the spatial derivative, $\mathbf{f}_x$, into axes which are photometrically variant and
photometrically invariant.

6.3.1   Spherical Color Spaces

The spherical color transformation, see Fig. 6.1a, is given by:

$$\left(\begin{array}{c} \theta \\ \varphi \\ r \end{array} \right) = \left(\begin{array}{c} \arctan(\frac{G}{R}) \\ \arcsin\left(\frac{\sqrt{R^2+G^2}}{\sqrt{R^2+G^2+B^2}} \right) \\ r = \sqrt{R^2 + G^2 + B^2} \end{array} \right). \tag{6.7}$$

The spatial derivatives are transformed to the spherical coordinate system by:

$$S\left(\mathbf{f}_x \right) = \mathbf{f}_x^s = \left(\begin{array}{c} r \sin \varphi\, \theta_x \\ r\varphi_x \\ r_x \end{array} \right) = \left(\begin{array}{c} \frac{G_x R - R_x G}{\sqrt{R^2+G^2}} \\ \frac{R_x RB + G_x GB - B_x(R^2+G^2)}{\sqrt{R^2+G^2}\sqrt{R^2+G^2+B^2}} \\ \frac{R_x R + G_x G + B_x B}{\sqrt{R^2+G^2+B^2}} \end{array} \right). \tag{6.8}$$

The scale factors follow from the Jacobian of the transformation. They ensure that
the norm of the derivative remains constant under transformation, hence $|\mathbf{f}_x| = |\mathbf{f}_x^s|$.
In the spherical coordinate system the derivative vector is a summation of a shadow-
shading variant part, $\mathbf{S}_x = (0, 0, r_x)^T$ and a shadow-shading quasi-invariant part,
given by $\mathbf{S}_x^c = (r \sin \varphi \theta_x, r\varphi_x, 0)^T$.

6.3.2   Opponent Color Spaces

The opponent color space, see Fig. 6.1b, is given by:

$$\left(\begin{array}{c} o1 \\ o2 \\ o3 \end{array} \right) = \left(\begin{array}{c} \frac{R-G}{\sqrt{2}} \\ \frac{R+G-2B}{\sqrt{6}} \\ \frac{R+G+B}{\sqrt{3}} \end{array} \right). \tag{6.9}$$

For this the following transformation of the derivatives follows:

$$O\left(\mathbf{f}_x \right) = \mathbf{f}_x^o = \left(\begin{array}{c} o1_x \\ o2_x \\ o3_x \end{array} \right) = \left(\begin{array}{c} \frac{1}{\sqrt{2}}\left(R_x - G_x \right) \\ \frac{1}{\sqrt{6}}\left(R_x + G_x - 2B_x \right) \\ \frac{1}{\sqrt{3}}\left(R_x + G_x + B_x \right) \end{array} \right). \tag{6.10}$$

The opponent color space decorrelates the derivative with respect to specular changes.
The derivative is divided into a specular variant part, $\mathbf{O}_x = (0, 0, o3_x)^T$, and a specular
quasi-invariant part $\mathbf{O}_x^c = (o1_x, o2_x, 0)^T$.
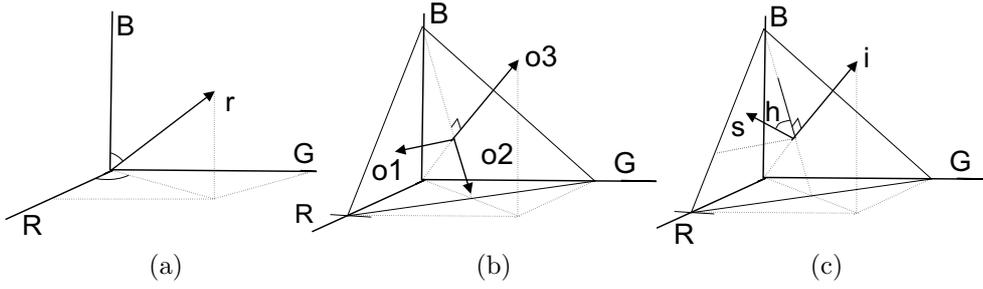
Figure 6.1: *The spherical, opponent and hue-saturation-intensity coordinate system.*

6.3.3   Hue-Saturation-Intensity Color Spaces

The well known hue-saturation-intensity is given by

$$
\begin{pmatrix} h \\ s \\ i \end{pmatrix} = \begin{pmatrix} \arctan\left(\frac{o1}{o2}\right) \\ \sqrt{o1^2 + o2^2} \\ o3 \end{pmatrix}.
\tag{6.11}
$$

The transformation of the spatial derivatives into the *hsi*-space decorrelates the derivative with respect to specular, shadow and shading variations,

$$
H\left(\mathbf{f}_x\right) = \mathbf{f}_x^h = \begin{pmatrix} s\,h_x \\ s_x \\ i_x \end{pmatrix} = \begin{pmatrix} \frac{(R(B_x-G_x)+G(R_x-B_x)+B(G_x-R_x))}{\sqrt{2(R^2+G^2+B^2-RG-RB-GB)}} \\ \frac{R(2R_x-G_x-B_x)+G(2G_x-R_x-B_x)+B(2B_x-R_x-G_x)}{\sqrt{6(R^2+G^2+B^2-RG-RB-GB)}} \\ \frac{(R_x+G_x+B_x)}{\sqrt{3}} \end{pmatrix}.
$$
$$
\tag{6.12}
$$

The shadow-shading-specular variant is given by $\mathbf{H}_x = (0, 0, i_x)^T$ and the shadow-shading-specular quasi-invariant by $\mathbf{H}_x^c = (sh_x, s_x, 0)^T$.

Since the length of a vector is not changed by coordinate transformations, the norm of the derivative remains the same in all three representations $|\mathbf{f}_x| = |\mathbf{f}_x^c| = |\mathbf{f}_x^o| = |\mathbf{f}_x^h|$. For both the opponent color space and the hue-saturation-intensity color space, the photometrically variant direction is given by the $L1$ norm of the intensity. For the spherical coordinate system the variant is equal to the $L2$ norm of the intensity.

We discussed three color spaces which decorrelate the color spaces with respect to various physical events. In the decorrelated color spaces often occurring physical variations, such as intensity changes, will only influence the photometric variant axes. In the next section the statistics of color image derivatives are examined in these decorrelated color spaces.

6.4   Statistics of Color Images

As discussed in Section 6.2 the information content of a descriptor depends on the probability of the derivatives, see Eq. 6.3 and Eq. 6.4. In this section we investigate

(a)   (b)   (c)
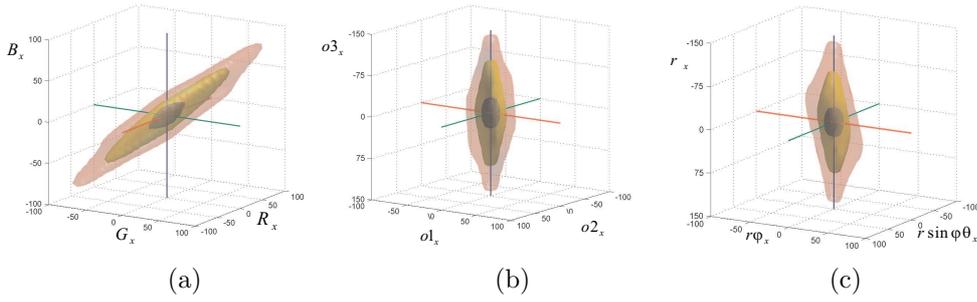
Figure 6.2: *The histograms of the distribution of the transformed derivatives of the Corel image database in respectively the (a) RGB coordinates, (b) the opponent coordinates and (c) the spherical coordinates. The three planes correspond with the isosalient surfaces which contain (from dark to light) respectively* 90%, 99%, $99.9\%t$ *of the total number of pixels (see also color plate C.17).*

the statistics of color derivatives in the decorrelated color spaces proposed in Section 6.3. From the statistics we aim to find a mathematical description of surfaces of equal probability, so called isosalient surfaces. Since a description of these surfaces leads to the solution of Eq. 6.5.

The statistics of color images are shown for the Corel database [15], which consists of 40,000 images after the exclusion of black and white images. In Fig. 6.2 the distributions of the first order derivatives, f_x, are given for the various color coordinate systems described in section 6.3 (HSI has been left out due to space considerations). The isosalient surfaces show a remarkably simple structure, approximately similar to an ellipsoid. For all three color spaces, the third coordinate coincides with the axis of maximum variation (i.e. the intensity). For the opponent and the spherical coordinate system, the first and second coordinate are rotated, with rotation matrix R^ϕ, so that the first coordinate coincides with the axis of minimum variation

$$\begin{aligned}
\left(r\sin\tilde{\varphi}\,\tilde{\theta}_x, r\tilde{\varphi}_x\right)^T &= R^\phi \left(r\sin\varphi\theta_x, r\varphi_x\right)^T \\
(\tilde{o}1_x, \tilde{o}2_x)^T &= R^\phi \left(o1_x, o2_x\right)^T.
\end{aligned}$$

(6.13)

The tilde indicates the color space transformation with the aligned axes. Similarly, the aligned transformations are given by $\tilde{S}(\mathbf{f}_x) = \mathbf{f}_x^{\tilde{s}}$ and $\tilde{O}(\mathbf{f}_x) = \mathbf{f}_x^{\tilde{o}}$.

After alignment of the axes isosalient surfaces of the derivative histograms can be approximated by ellipsoids

$$\left(\alpha h_x^1\right)^2 + \left(\beta h_x^2\right)^2 + \left(\gamma h_x^3\right)^2 = R^2$$

(6.14)

where $\mathbf{h}_x = h(\mathbf{f}_x) = \left(h_x^1, h_x^2, h_x^3\right)^T$ and h is one of the transformations $\tilde{S}$, $\tilde{O}$, or H.

6.5   Boosting Color Saliency

We now return to our goal, that is to incorporate color distinctiveness into salient point detection. Or mathematically, to find the transformation for which vectors with equal information content have equal impact on the saliency function. In the previous section we saw that derivatives of equal saliency form an ellipsoid. Since Eq. 6.14 is equal to

$$\left(\alpha h_x^1\right)^2 + \left(\beta h_x^2\right)^2 + \left(\gamma h_x^3\right)^2 = |\mathbf{\Lambda} h\left(\mathbf{f}_x\right)|^2 \tag{6.15}$$

the following holds

$$p\left(\mathbf{f}_x\right) = p\left(\mathbf{f}_x'\right) \leftrightarrow |\mathbf{\Lambda} h\left(\mathbf{f}_x\right)| = \left|\mathbf{\Lambda}^T h\left(\mathbf{f}_x'\right)\right|, \tag{6.16}$$

where $\mathbf{\Lambda}$ is a 3x3 diagonal matrix with $\mathbf{\Lambda}_{11} = \alpha$, $\mathbf{\Lambda}_{22} = \beta$, and $\mathbf{\Lambda}_{33} = \gamma$. Λ is restricted to $\Lambda_{11}^2 + \Lambda_{22}^2 + \Lambda_{33}^2 = 1$. The desired saliency boosting function (see Eq. 6.5) is obtained

$$g\left(\mathbf{f}_x\right) = \mathbf{\Lambda} h\left(\mathbf{f}_x\right). \tag{6.17}$$

By a rotation of the color axes followed by a rescaling of the axis, the oriented isosalient ellipsoids are transformed into spheres, and thus vectors of equal saliency are transformed into vectors of equal length.

6.5.1   Influence of Color Saliency Boosting on Repeatability

In the introduction two criteria for salient point detection were described, namely distinctiveness and repeatability. The color boosting algorithm is designed to focus on color distinctiveness, while adopting the geometrical characteristics of the operator to which it is applied. In this section we examine the influence of color boosting on the repeatability. We identify two phenomena which influence the repeatability of $g\left(\mathbf{f}_x\right)$. Firstly, by boosting the color saliency an anisotropic transformation is carried out. This will reduce the signal-to-noise ratio negatively. Secondly, by boosting the photometric invariant directions more than the photometric variant directions, we improve robustness with respect to scene accidental changes.

For isotropic uncorrelated noise, ε, the measured derivative $\hat{\mathbf{f}}_x$ can be written as

$$\hat{\mathbf{f}}_x = \mathbf{f}_x + \varepsilon \tag{6.18}$$

and after color saliency boosting

$$g\left(\hat{\mathbf{f}}_x\right) = g\left(\mathbf{f}_x\right) + \mathbf{\Lambda}\varepsilon. \tag{6.19}$$

Note that isotropic noise remains unchanged under the orthogonal curvilinear transformations. Assume the worst case in which $\mathbf{f}_x$ only has signal in the photometric variant direction, then the noise can be written as

$$\frac{|g\left(\mathbf{f}_x\right)|}{|\mathbf{\Lambda}\varepsilon|} \approx \frac{\Lambda_{33}\,|\mathbf{f}_x|}{\Lambda_{11}\,|\varepsilon|}. \tag{6.20}$$

	$\mathbf{f}_x$	$\|\mathbf{f}_x\|_1$	$\mathbf{f}_x^{\tilde{s}}$	$\tilde{\mathbf{S}}_x^c$	$\mathbf{f}_x^{\tilde{o}}$	$\tilde{\mathbf{O}}_x^c$	$\mathbf{f}_x^h$	$\mathbf{H}_x^c$
Λ_{11}	0.577	1	0.851	0.856	0.850	0.851	0.858	1
Λ_{22}	0.577	-	0.515	0.518	0.524	0.525	0.509	0
Λ_{33}	0.577	-	0.099	0	0.065	0	0.066	0

Table 6.1: *The diagonal entries of* Λ *for the Corel data set computed for Gaussian derivatives with* $\sigma = 1$.

Hence, the signal-to-noise ratio reduces by $\frac{\Lambda_{11}}{\Lambda_{33}}$, which will negatively influence repeatability to geometrical and photometrical changes.

The second phenomena which influences repeatability is the gain in photometric robustness. By boosting color saliency the influence of the photometric variant direction diminishes while the influence of the quasi-invariant directions increases. As a consequence the repeatability under photometric changes, such as changing illumination and viewpoint, increases.

Depending on the task at hand, distinctiveness may be less desired than signal-to-noise. For this purpose the α parameter is proposed, which allows for choosing between best signal-to-noise characteristics, $\alpha = 0$, and best information content, $\alpha = 1$:

$$g^\alpha\left(\mathbf{f}_x\right) = \alpha\mathbf{\Lambda}h\left(\mathbf{f}_x\right) + \left(1 - \alpha\right)h\left(\mathbf{f}_x\right). \qquad (6.21)$$

For $\alpha = 0$ this is equal to color gradient-based salient point detection.

6.6   Experiments and Illustrations

Color saliency boosting is tested on: information content and repeatability. The salient points based on color saliency boosting are compared to luminance, *RGB* gradient, and the quasi-invariant-based salient point detectors. The generality of the approach is illustrated by applying color boosting to several existing feature detectors.

6.6.1   Initialization

Experiments are performed on a subset of 1000 randomly chosen images from the Corel data set. Before color saliency boosting can be applied, the Λ-parameters (Eq.6.15) have to be initialized by fitting ellipses to the histogram of the data set. First the axes of the opponent and the spherical transformation are aligned by Eq. 6.13. Next, the axes of the ellipsoid are derived by fitting the isosaliency surface which contains 99 percent of the pixels of the histogram of the Corel data set. The results for the various transformations are summarized in Table 6.1. The relation between the axes in the various color spaces clearly confirms the dominance of the luminance axis in the *RGB*-cube, since Λ_{33}, the multiplication-factor of the luminance axis, is much smaller than the color-axes multiplication factors, Λ_{11} and Λ_{22}.

To give an idea on how the Λ-parameters change when changing the data set, we also estimated the Λ parameters for two other data sets, the Soil data [42] and a table-tennis sequence (see Fig. 6.3a,c). For the Soil data and the opponent color model the
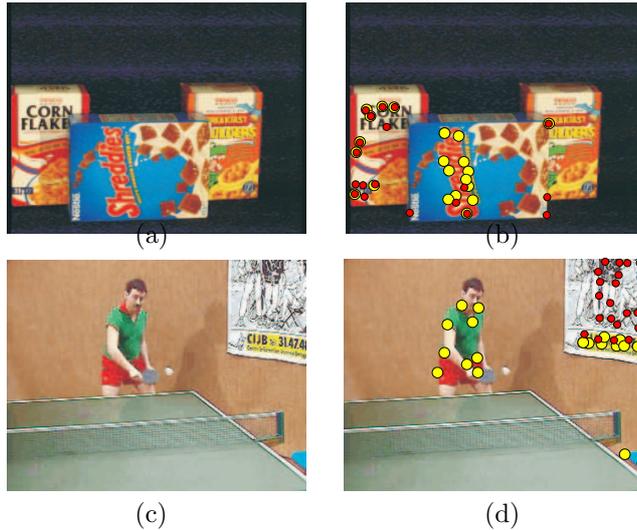
Figure 6.3: *(a) Example Soil data set and (c) frame from table-tennis sequence. (b) and (d) results of Harris detector (red dots) and the Harris detector with color boosting (yellow dots). The red dots mainly coincide with black and white events, while the yellow dots are focussed on colorful points (see also color plate C.18).*

Λ-parameters are $\Lambda_{11} = 0.542$, $\Lambda_{22} = 0.780$, and $\Lambda_{33} = 0.313$. Since this set consists of colorful objects the luminance axis is less suppressed than for the Corel set. For the tennis sequence the difference with Corel is smaller, $\Lambda_{11} = 0.588$, $\Lambda_{22} = 0.799$, and $\Lambda_{33} = 0.124$. A change in Λ-parameters can have various causes such as the quality of the camera, the applied compression and the different content of the data.

We have chosen the Harris point detector [27] to test color boosting in experiment B, C, and D. It is computed with

$$H^{\sigma}\left(\mathbf{f}_x, \mathbf{f}_y\right) = \overline{\mathbf{f}_x \cdot \mathbf{f}_x \, \mathbf{f}_y \cdot \mathbf{f}_y} - \overline{\mathbf{f}_x \cdot \mathbf{f}_y}^2 - k\left(\overline{\mathbf{f}_x \cdot \mathbf{f}_x} + \overline{\mathbf{f}_y \cdot \mathbf{f}_y}\right)^2 \qquad (6.22)$$

by substituting $\mathbf{f}_x$ and $\mathbf{f}_y$ by $g\left(\mathbf{f}_x\right)$ and $g\left(\mathbf{f}_y\right)$. The bar $\bar{.}$ indicates convolution with a Gaussian filter and the dot indicates the inner product. We applied Gaussian derivatives of $\sigma = 1$ and Gaussian smoothing with $\sigma = 3$.

6.6.2 Color Distinctiveness

Here we examine if color boosting improves the color distinctiveness of the Harris detector. In [64], the Harris detector has already been shown to outperform other detectors both on 'shape' distinctiveness and repeatability. The color distinctiveness of salient point detectors is described by the information content of the descriptors extracted at the locations of the salient points. From the combination of Eq. 6.3 and Eq. 6.4, it follows that the total information is computed by summing the information

	standard descriptor						normalized descriptor					
	20 points			100 points			20 points			100 points		
method	inf.	incr(%)	decr(%)	inf.	incr.	decr.	inf.	incr.	decr.	inf.	incr.	decr.
$\mathbf{f}_x$	20.4	-	-	20.0	-	-	13.2	-	-	13.9	-	-
$|\mathbf{f}_x|_1$	19.9	0	1.4	19.8	0	0.8	13.0	0	2.7	13.8	0	1.0
$\tilde{\mathbf{S}}_x^c$	22.2	45.5	10.1	20.4	9.1	17.7	17.9	<u>92.9</u>	0.9	16.2	<u>69.8</u>	2.8
$\mathbf{f}_x^{\tilde{s}}$	22.3	49.4	.6	20.8	13.1	1.3	16.9	86.9	0.6	15.5	57.6	.7
$\tilde{\mathbf{O}}_x^c$	22.6	51.4	12.9	20.5	12.0	34.2	<u>18.9</u>	92.5	1.3	<u>16.5</u>	64.6	10.8
$\mathbf{f}_x^{\tilde{o}}$	<u>23.2</u>	<u>62.6</u>	0.0	<u>21.4</u>	<u>21.5</u>	0.9	18.4	88.2	0.3	16.4	65.0	1.7
$\mathbf{H}_x^c$	21.0	21.7	43.4	19.0	1.8	77.4	17.3	77.1	10.9	14.8	31.7	37.9
$\mathbf{f}_x^h$	23.0	57.2	0.3	21.3	16.7	1.1	18.3	87.4	0.5	16.2	62.3	2.2
rand.	14.4	0	99.8	14.4	0	100	10.1	2.7	89.1	10.2	.6	96.7

Table 6.2: *The information content of salient point detectors. Measured in 1. information content and 2. the percentage of images for which a substantial decrease (-5%) or increase ($+5\%$) of the information content occurs. The experiment is performed with both 20 or 100 salient points per image. The experiment is repeated with a normalized descriptor which is invariant for luminance changes.*

of the zeroth and first order part, $I(\mathbf{v}) = I(\mathbf{f}) + I(\mathbf{f}_x) + I(\mathbf{f}_y)$. The information content of the parts is computed from normalized histograms with

$$I(\mathbf{f}) = -\sum_i p_i \log(p_i) \qquad (6.23)$$

where p_i are the probabilities of the bins of the histogram of $\mathbf{f}$.

The results for 20 and 100 salient points per image are shown in Table 6.2. Next to the absolute information content we have also computed the relative information gain with respect to the information content of the color gradient based Harris detector. For this purpose the information content on a single image is defined as

$$I = -\sum_{j=1}^{n} \log(p(v_j)) \qquad (6.24)$$

where $j = 1, 2, ...n$ and n is the number of salient points in the image. Here $p(v_j)$ is computed from the global histograms, which allows comparison of the results per image. The information content change is considered substantially for a 5 percent increase or decrease.

The highest information content is obtained with $\mathbf{f}_x^{\tilde{o}}$, which is the color saliency boosted version of the opponent derivatives. The boosting results in an increase of 7% to 13% of the information content compared to the color gradient based detector. On the images of the Corel set this resulted in a substantial increase on 22% to 63% of the images. The advantage of color boosting diminishes when increasing the number of salient points per image. This is caused by the limited number of color clues
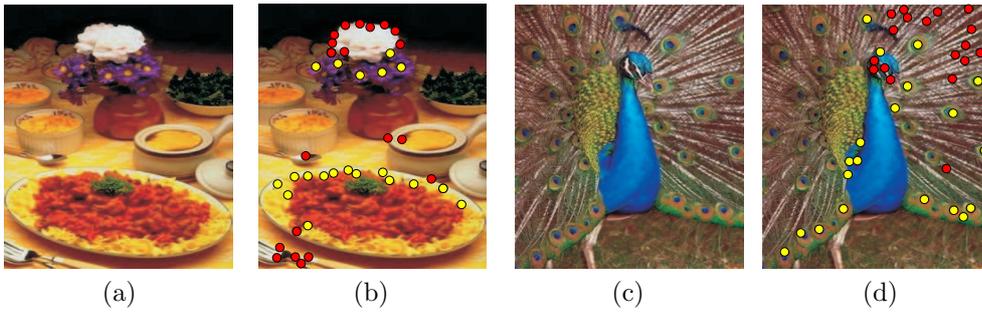
Figure 6.4: *(a) and (c) Corel input images. (b) and (d) results of Harris detector (red dots) and the Harris detector with color boosting (yellow dots). The red dots mainly coincide with black and white events, while the yellow dots are focussed on colorful points (see also color plate C.19).*

in many of the images, which is especially visible for the results of the photometric quasi-invariants, $\tilde{\mathbf{S}}_x^c$, $\tilde{\mathbf{O}}_x^c$, or $\mathbf{H}_x^c$. These detectors discard all intensity information, which in the case of 100 salient points per image results in many images with a substantial decrease in information content. Finally, it is noteworthy to see how small the difference is between luminance and RGB-based Harris detection. Since the intensity direction also dominates the RGB derivatives, using RGB-gradient instead of luminance-based Harris detection only results in a substantial increase in information content in 1% of the images.

It might be desirable for the descriptor to be invariant for scene incidental events like shading and shadows [63]. In these cases the information content of the normalized descriptor, which is invariant to luminance changes, better reflects the information content of the salient point detector

$$v = \left(\frac{R}{|\mathbf{f}|}, \frac{G}{|\mathbf{f}|}, \frac{B}{|\mathbf{f}|}, \frac{R_x}{|\mathbf{f}_x|}, \frac{G_x}{|\mathbf{f}_x|}, \frac{B_x}{|\mathbf{f}_x|}, \frac{R_y}{|\mathbf{f}_y|}, \frac{G_y}{|\mathbf{f}_y|}, \frac{B_y}{|\mathbf{f}_y|} \right). \tag{6.25}$$

The results of the normalized descriptor are given in the right half of Table 6.2. The increase in information content of the quasi-invariants and the color boosted detectors stands out even more, with substantial gains in information content of up to 90%. Here the quasi-invariants based detectors outperform the other detectors.

In Fig. 6.4 results of the RGB-gradient based and color boosted Harris detector are depicted. From a color information point of view, the RGB-gradient based method does a poor job. Most of the salient points have a black and white local neighborhood, with a low color saliency. The salient points after color boosting focus on more distinctive points. Similar results are depicted in Fig. 6.3b,d, where the results are shown computed with the Λ-parameters belonging to the data sets of these images.

6.6.3 Repeatability: signal-to-noise

Repeatability measures the stability with respect to varying viewing conditions. As indicated in section 6.5.1 color saliency boosting reduces the signal-to-noise ratio.

Repeatability with respect to geometrical changes, scaling, and affine transformations are considered a property of the detector and will not be considered here.

The loss of repeatability caused by color saliency boosting is examined by adding uniform, uncorrelated Gaussian noise of $\sigma = 10$. This yields a good indication of loss in signal-to-noise, which in its turn will influence results of repeatability under other variations, such as zooming, illumination changes, and geometrical changes. Repeatability is measured by comparing the Harris points detected in the noisy image to the points in the noise-free images. The results in Fig. 6.5a correspond to the expectation made by Eq. 6.20, namely the larger the difference between Λ_{11} and Λ_{33}, the poorer the repeatability.

In Fig. 6.5b the information content and repeatability as a function of the color boosting, determined by the α-parameter, is given (see Eq.6.21). The experiment is performed by applying color boosting to the opponent color space. The results show that information content increases at the cost of stability. Depending on the application a choice should be made about the amount of color saliency boosting.

6.6.4 Repeatability: photometric variation

Photometric robustness increases with color boosting, as discussed in Section 6.5.1. In Fig. 6.6 the dependance of repeatability is tested on two image sequences with changing illumination conditions [53]. The experiment was performed by applying color boosting to the spherical color space, since changes due to shadow-shading will be along the photometric variant direction of the spherical system. For these experiments two intertwining phenomena can be observed: the improved photometric invariance and the deterioration of signal-to-noise ratio with increasing α. For the nuts-sequence, with very prominent shadows and shading, the photometric invariance is dominant, while for the fruit-basket the gained photometric invariance only improves performance slightly for medium α values. For total color saliency boosting, $\alpha = 1$ the loss of repeatability, due to loss of signal-to-noise, is substantial.

method	20 points	100 points
$\mathbf{f}_x$	88	84
$|\mathbf{f}_x|_1$	88	83
$\tilde{\mathbf{S}}_x^c$	53	42
$\mathbf{f}_x^{\tilde{s}}$	62	54
$\tilde{\mathbf{O}}_x^c$	46	34
$\mathbf{f}_x^{\tilde{o}}$	51	41
$\mathbf{H}_x^c$	35	25
$\mathbf{f}_x^h$	52	42
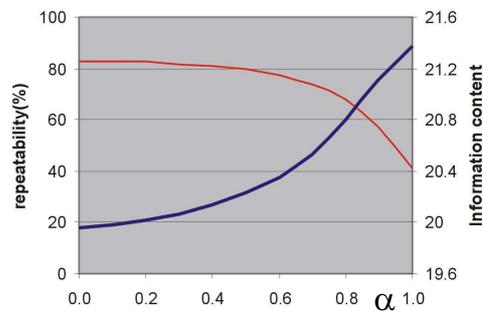
(a)   (b)

Figure 6.5: *(a) The percentage of Harris points which remain detected after adding Gaussian uncorrelated noise. (b) The information content (blue line) and the repeatability (red line) as a function of the amount of color saliency boosting.*

Figure 6.6: *(a),(b) Two frames from two sequences with changing illumination conditions. (c) Repeatability as a function of the amount of color saliency boosting for the two sequences. Dotted line for the nuts-sequence and the continuous line for the fruit-basket sequence (see also color plate C.20).*

6.6.5 Illustrations Generality

Color saliency boosting can be applied on all functions which can be written as a function of the local derivatives. Here we apply it to three different feature detectors. First we apply saliency boosting to the focus point detector which was originally proposed by Reisfeld et al. [60] and recently extended to color by Heidemann [29]. The detector focuses on the center of locally symmetric structures. Fig. 6.7b shows the saliency map as proposed in [29]. In Fig. 6.7c the result after saliency boosting is depicted. Although focus point detection is already an extension from luminance to color, black-and-white transition still dominate the result. Only after boosting the color saliency, the less interesting black-and-white structures in the image are ignored and most of the red Chinese signs are found. Similar difference in performance is obtained by applying color boosting to the linear symmetry detector proposed by Bigün [6]. This detector focuses on corner and junction like structures. The *RGB* gradient based method focuses mainly on black-and-white events while the more salient signboards are found only after color saliency boosting.

As a final illustration we illustrate that color saliency boosting can easily be applied to gradient based methods. In third row of Fig. 6.7 color boosting is applied to a gradient based segmentation algorithm proposed by Jermyn and Ishikawa [36]. The algorithm finds globally optimal regions and boundaries. In Fig. 6.7b and c respectively the *RGB* gradient and the color boosted gradient are depicted. While the *RGB*-gradient based segmentation is distracted by the many black-and-white events in the background, the color boosted segmentation finds the salient traffic signs.

6.7   Conclusions

In this chapter color distinctiveness is explicitly incorporated in the design of salient point detectors. The method, called color saliency boosting, can be incorporated into

(a)	(e)	(i)
(b)	(f)	(j)
(c)	(g)	(k)
(d)	(h)	(l)
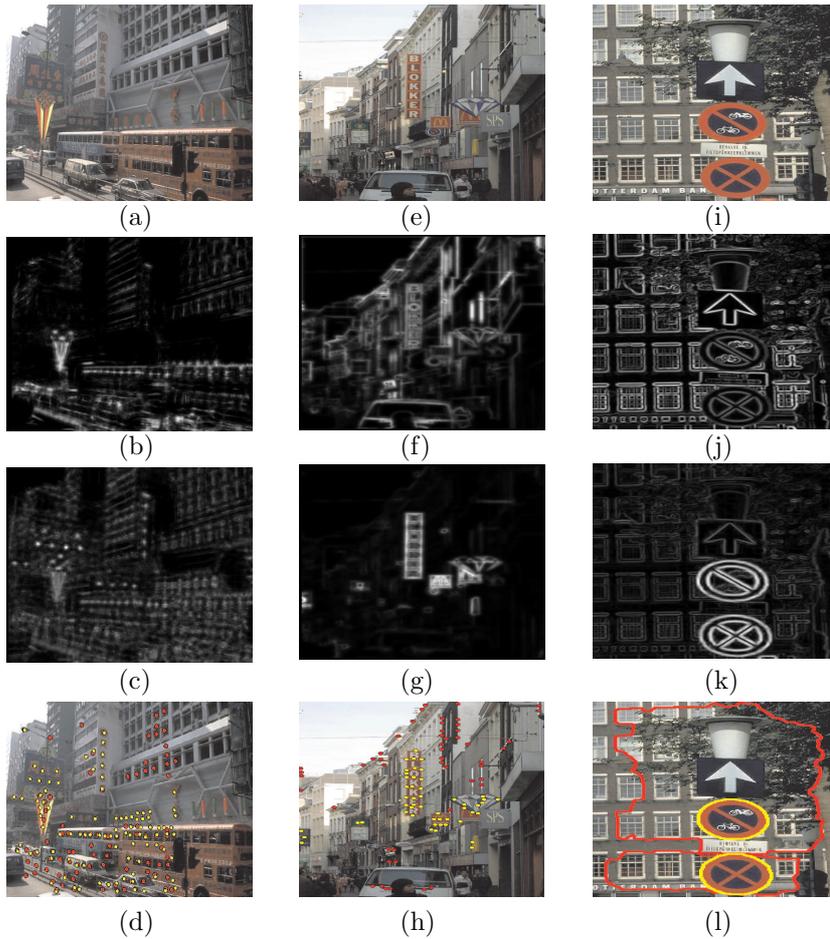
Figure 6.7: *Respectively, input image, RGB-gradient based saliency map, color boosted saliency map and the results with red dots (lines) for gradient-based method and yellow dots (lines) for salient points after color saliency boosting. Results (a),(b),(c),(d) for the focus points, (e),(f),(g),(h) for the symmetry points and (i),(j),(k),(l) for the global optimal regions and boundary method (see also color plate C.21).*

existing detectors which are mostly focused on shape distinctiveness. Saliency boosting is based upon an analysis of the statistics of color image derivatives. Isosalient derivatives form ellipsoids in the color derivative histograms. This fact is exploited to adapt derivatives in such a way that equal saliency implies equal impact on the saliency map. Experiments show that color saliency boosting substantially increases the information content of the detected points. A substantial information content increase is obtained on up to $20 - 60\%$ of the Corel images. Further, the generality of the method is illustrated by applying color boosting to various point detectors.

Chapter 7

Summary and Conclusions

7.1   Summary

In this thesis, we explore methods to exploit the extra information available in color images as opposed to grey-value images. We indicate two main advantages of using color data over luminance data. Firstly, color data contains a richer photometric description of the local structure from which various causes for variations in the image can be distinguished. This richer description allows for example to separate highly informative object edges from less informative shadow edges. A second advantage of color over luminance is that color can be used to improve the distinctiveness of salient point detectors. An analysis of the distribution of colors in the world allows to distinguish between low frequent and therefore highly informative colors, and high frequent and therefore less informative colors. The observation of color distinctiveness can be incorporated in existing salient point detectors. The two above mentioned advantages let to the following main objectives for this thesis:

1. **From Luminance to Color:** Extend luminance-based algorithms to color in a mathematically sound way. One consequence is that color image enhancement methods do not introduce new chromaticities. A second implication is that for differential-based algorithms the derivatives of the separate channels should be combined without loss of derivative information.

2. **Photometric Information:** Compute photometric invariant differential information in a robust way. We focus on the class of applications for which no a-priori knowledge of the noise characteristics of the acquisition system is available.

3. **Color Distinctiveness:** Improve the distinctiveness of salient point detection algorithms by explicitly incorporating color statistics into the detector design.

The results obtained in the thesis are discussed per chapter in the following paragraphs:

Chapter 2: Least Squares and Robust Estimation of Local Image Structure. In this chapter we propose the Gaussian facet model, as a generalization of the classic Haralick facet model, which constructs a polynomial approximation of the unsmoothed image. The measured differential structure therefore is closer to the 'real' structure than the differential structure measured using Gaussian derivatives. At the points in an image where the differential structure changes abruptly (because of discontinuities in the imaging conditions, e.g. a material change, or a depth discontinuity) both the Gaussian derivatives and the Gaussian facet model diffuse the information from both sides of the discontinuity (smoothing across the edge). Robust estimators that are classically meant to deal with statistical outliers can also be used to deal with these 'mixed model distributions'. In this paper we introduce the robust estimators of local image structure. We start with the Gaussian facet model where we replace the quadratic error norm with a robust (Gaussian) error norm, which leads to a robust Gaussian facet model. Examples are given for luminance and color images, and for both zero and higher order differential structure.

Chapter 3: Edge and Corner Detection by Photometric Quasi-Invariants. We propose a new class of derivatives which we refer to as quasi-invariants. These quasi-invariants are derivatives which share with full photometric invariants the property that they are insensitive for certain photometric edges, such as shadows or specular edges, but without the inherent instabilities of full photometric invariants. Experiments show that the quasi-invariant derivatives are less sensitive to noise and introduce less edge displacement than full invariant derivatives. Moreover, quasi-invariants significantly outperform the full invariant derivatives in terms of discriminative power.

Chapter 4: Curvature Estimation in Oriented Patterns Using Curvilinear Models. Curved oriented patterns are dominated by high frequencies and exhibit zero gradients on ridges and valleys. Existing curvature estimators fail here. The characterization of curved oriented patterns based on translation invariance lacks an estimation of local curvature and yields a biased curvature-dependent confidence measure. In chapter 4, we use parameterized curvilinear models to measure the amount of local gradient energy along the model gradient as a function of model curvature. Minimizing the residual energy yields a closed-form solution for the local curvature estimate and the corresponding confidence measure. We show that simple curvilinear models are applicable in the analysis of a wide variety of curved oriented patterns.

Chapter 5: Robust Photometric Invariant Features from the Color Tensor. In this chapter we focus on the structure tensor, or color tensor, which adequately handles the vector nature of color images. Further, we combine the features based on the color tensor with photometric invariant derivatives to arrive at photometric invariant features. We circumvent the drawback of unstable photometric invariants by deriving an uncertainty measure to accompany the photometric invariant derivatives. The uncertainty is incorporated in the color tensor, hereby allowing the computation of robust photometric invariant features. The combination of the photometric invariance theory and tensor-based features allows for detection of a variety of features such as photometric invariant edges, corners, optical flow and curvature. The proposed features are tested for noise characteristics and robustness to photometric changes. Experiments show that the proposed features are robust to scene incidental

events and that the proposed uncertainty measure improves the applicability of full invariants.

Chapter 6: Boosting Color Saliency in Image Feature Detection. In this chapter color distinctiveness is explicitly incorporated into the design of saliency detection. The algorithm, called color saliency boosting, is based on an analysis of the statistics of color image derivatives. Isosalient color derivatives are shown to form ellipsoidal surfaces. Based on this remarkable statistical finding, isosalient derivatives are transformed by color boosting to have equal impact on the saliency. Color saliency boosting is designed as a generic method easily adaptable to existing, mostly shape distinctiveness focussed, feature detectors. Results show that substantial improvements in information content are acquired by targeting color salient features. Further, the generality of the method is illustrated by applying color boosting to multiple existing saliency methods.

7.2   Conclusions

In this thesis we have proposed theory and techniques to augment the usefulness of color for computer vision. The first objective of the thesis is to extend luminance-based algorithms to color in a mathematical sound way. For color image filtering this implies that the correlation between the channels has to be taken into account to prevent the introduction of undesired new chromaticities. To this end, a robust estimator of local image structure is proposed. The estimation is based on the iterative use of a spatial-tonal Gaussian filter which is based on both the spatial distance and the tonal distance between pixel values. Application of the robust estimator results in efficient noise reduction with only little loss of contrast, and without the introduction of new chromaticities. For differential-based algorithms the extension from luminance to color poses a different problem. The problem is how to combine the differential information of the separate channels. We observe that tensor mathematics solves this problem. Therefore we give an overview of tensor-based features and we show how to extend them to color.

The second objective of the thesis is to design robust photometric invariant differential operators. We distinguish between feature detection, i.e. the localization of a feature, and feature extraction, i.e. the extraction of a descriptor of a local neighborhood at a certain location in the image. Firstly, for feature detection, a set of derivative filters is proposed which are coined quasi-invariants. These filters outperform existing full photometric invariant derivatives in terms of discriminative power and localization. Secondly, for feature extraction, we derive an uncertainty measure to accompany full-invariant derivatives. Color features which incorporate this uncertainty measure are shown to outperform existing full invariant features. The proposed color features include: edges, corners, symmetry points, circle detectors, and optical flow.

The third objective of the thesis is to improve the distinctiveness of salient point detection algorithms by explicitly incorporating color statistics into the detector design. From information theory it is known that rare events, i.e. events with a low

frequency of occurrence, have high information content. Salient point detection aims at detecting salient, and hence highly informative points in the image. Most existing salient point detectors are luminance-based and are computed from the differential structure of the image. An analysis of the statistics of color derivatives for a large data set of real world pictures reveals a remarkable phenomenon; derivatives with equal frequency, and hence equal information content, form ellipsoid surfaces in derivative space. We exploit this phenomenon by adjusting the saliency functions in such a way that points with equal information content have equal influence on the saliency function. This process is called color saliency boosting, and it has been proven to substantially increase the information content of the detected salient points.

Bibliography

[1] J. Astola, P. Haavisto, and Y. Neuvo. Vector median filters. *IEEE Proceedings*, 78(4):678–689, April 1990.

[2] P. Bakker, L.J. van Vliet, and P.W. Verbeek. Edge preserving orientation adaptive filtering. In M. Boasson, J.A. Kaandorp, J.F.M. Tonino, and M.G. Vosselman, editors, *ASCI'99, Proc. 5th Annual Conference of the Advanced School for Computing and Imaging*, pages 207–213, 1999.

[3] D. H. Ballard. Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition*, 12(2):111–122, 1981.

[4] J. Barron and R. Klette. Quantitative color optical flow. In *Int. Conf. on Pattern Recognition*, pages 251–255, Vancouver, Canada, 2002.

[5] P.J. Besl, J.B. Birch, and L.T. Watson. Robust window operators. *Machine Vision and Applications*, 2:179–191, 1989.

[6] J. Bigun. Pattern recognition in images by symmetry and coordinate transformations. *Computer Vision and Image Understanding*, 68(3):290–307, 1997.

[7] J. Bigun and M.H. du Buf. Symmetry interpretation of complex moments and the local power spectrum. *J. Visual Comm. and Image Representation*, 6:154–163, 1995.

[8] J. Bigun, G. Granlund, and J. Wiklund. Multidimensional orientation estimation with applications to texture analysis and opitcal flow. *IEEE trans. on pattern analysis and machine intelligence*, 13(8):775–790, 1991.

[9] M. Black, G. Sapiro, D. Marimont, and D. Heeger. Robust anisotropic diffusion. *IEEE Trans. Image Processing*, 7(3):421–432, March 1998.

[10] J. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8(6):679–698, 1986.

[11] Silvano Di Zenzo. Note: A note on the gradient of a multi-image. *Computer Vision, Graphics, and Image Processing*, 33(1):116–125, 1986.

[12] G. Farnebäck. *Spatial Domain Methods for orientation and velocity measurements*. PhD thesis, Linköping University, 1999.

[13] R. Fergus, P. Perona, and A. Zisserman. Object class recognition by unsupervised scale-invariant learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 264–271, June 2003.

[14] L. M. J. Florack, B. M. ter Haar Romeny, J. J. Koenderink, and M. A. Viergever. Scale and the differential structure of images. *Image and Vision Computing*, 10(6):376–388, July/August 1992.

[15] Corel Gallery. www.corel.com.

[16] J.M. Geusebroek, R. van den Boomgaard, A.W.M. Smeulders, and H. Geerts. Color invariance. *IEEE Trans. Pattern Analysis Machine Intell.*, 23(12):1338–1350, 2001.

[17] Th. Gevers. *Color image invariant segmentation and retrieval*. PhD thesis, University of Amsterdam, 1997.

[18] Th. Gevers and A. Smeulders. Color based object recognition. *Pattern Recognition*, 32:453–464, March 1999.

[19] Th. Gevers and H.M.G. Stokman. Classification of color edges in video into shadow-geometry, highlight, or material transitions. *IEEE Trans. on Multimedia*, 5(2):237–243, 2003.

[20] Th. Gevers and H.M.G. Stokman. Robust histogram construction from color invariants for object recognition. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 26:113–118, 2004.

[21] P. Golland and A. M. Bruckstein. Motion from color. *Computer Vision and Image Understanding*, 68(3):346–362, December 1997.

[22] L.D. Griffin. Scale-imprecision space. *Image and Vision Computing*, 15:369–398, 1997.

[23] L. Haglund. *Adaptive Multi-Dimensional Filtering*. PhD thesis, Linkoping University, 1992.

[24] O. Hansen and J. Bigun. Local symmetry modeling in multi-dimensional images. *pattern Recognition Letters*, 13:253–262, 1992.

[25] R. M. Haralick and L. Watson. A facet model for image data. *Computer Graphics and Image Processing*, 15:113–129, 1981.

[26] R.M. Haralick and L.G. Shapiro. *Computer and Robot Vision, volume II*. Addison-Wesley, 1992.

[27] C. Harris and M. Stephens. A combined corner and edge detector. In *Proc. 4th Alvey Vision Conf.*, volume 15, pages 147–151, 1988.

[28] G. Healey. Segmenting images using normalized color. *IEEE Trans. Syst., Man, Cybern.*, 22:64–73, 1992.

[29] G. Heidemann. Focus-of-attention from local color symmetries. *IEEE trans. on pattern analysis and machine intelligence*, 26(7):817–847, 2004.

[30] M.H. Heuckel. An operator which locates edges in digital pictures. *J. Association for computing machinery*, 18:113–125, 1971.

[31] B. K. P. Horn and B. G. Schunk. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

[32] P.J. Huber. *Robust Statistics*. Probability and mathematical statistics. Wiley, 1981.

[33] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11):1254–1259, Nov 1998.

[34] L. Itti, C. Koch, and E. Niebur. Computation modeling of visual attention. *Nature Reviews Neuroscience*, 2(11):194–203, March 2001.

[35] A. Jain, L. Hong, and R. Bolle. On-line fingerprint verification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19:302–314, 1997.

[36] I.H. Jermyn and H. Ishikawa. Globally optimal regions and boundaries as minimum ratio weight cycles. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(10):1075–1088, Oct 2001.

[37] M. Kass and A. Witkin. Analyzing oriented patterns. *Computer Vision, Graphics, and Image Processing*, 37:362–385, 1987.

[38] G.J. Klinker and S.A. Shafer. A physical approach to color image understanding. *Int. Journal of Computer Vision*, 4:7–38, 1990.

[39] H. Knutsson and C.F. Westin. Normalised and differential convolution-methods for interpolation and filtering of incomplete and uncertain data. In *CVPR*, pages 515–523, 93.

[40] J J Koenderink and A J van Doom. Representation of local geometry in the visual system. *Biol. Cybern.*, 55(6):367–375, 1987.

[41] J.J. Koenderink and A.J. van Doorn. The structure of locally orderless images. *International Journal of Computer Vision*, 31(2/3):159–168, 1999.

[42] D. Koubaroulis, J. Matas, and J. Kittler. Evaluating colour-based object recognition algorithms using the soil-47 database. In *Asian Conference on Computer Vision*, 2002.

[43] M. Kuwahara, K. Hachimura, S. Eiho, and M. Kinoshita. Processing of ri-angiocardiographic images, 1976.

[44] M.S. Lee and G. Medioni. Grouping into regions, curves, and junctions. *Computer Vision Image Understanding*, 76(1):54–69, 1999.

[45] T. Lindeberg. Feature detection with automatic scale selection. *Int. J. Computer Vision*, 30(2):77–116, 1998.

[46] T. Lindeberg and J. Garding. Shape from texture from a multi-scale perspective. In H.-H. Nagel, editor, *Proc. 4th Int. Conf. on Computer Vision*, pages 683–691, 1993.

[47] D.G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. J. Computer Vision*, 60(2):91–110, 2004.

[48] B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proc. DARPA Image Understanding Workshop*, pages 121–130, 1981.

[49] D. Maio and D. Maltoni. Direct gray-scale munitiae detection in fingerprints. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 19:27–40, 1997.

[50] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proc. 8th Int'l Conf. Computer Vision*, volume 2, pages 416–423, July 2001.

[51] B.A. Maxwell and S.A. Shafer. Physics-based segmentation of complex objects using multiple hypothesis of image formation. *Computer Vision and Image Understanding*, 65:265–295, 1997.

[52] B.A. Maxwell and S.A. Shafer. Segmentation and interpretation of multicolored objects with highlights. *Computer Vision and Image Understanding*, 77:1–24, 2000.

[53] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. In *International Conference on Computer Vision and Pattern Recognition*, Madison, USA, 2003.

[54] K. Mikolajczyk and Cordelia Schmid. Scale and affine invariant interest point detectors. *Int. J. Computer Vision*, 60(1):62–86, 2004.

[55] P. Montesinos, V. Gouet, and R. Deriche. Differential Invariants for Color Images. In *Proceedings of 14th International Conference on Pattern Recognition*, pages 838–840, Brisbane, Australia, 1998.

[56] M. Nagao and T. Matsuyama. Edge preserving smoothing, 1979.

[57] Y. Ohta, T. Kanade, and T. Sakai. Color information for region segmentation. *Computer Graphics and Image Processing*, 13:222–241, 1980.

[58] PANTONE. edition 1992-1993 groupe basf, paris, france. pantone is a trademark of pantone inc.

[59] I. Pitas and A. N. Venetsanopoulos. Order statistics in digital image processing. *Proc. IEEE*, 80(12):1893–1921, 1992.

[60] D. Reisfeld, H. Wolfson, and Y. Yeshurun. Context free attentional operators: the generalized symmetry. *International Journal of Computer Vision*, 14:119–130, 1995.

[61] O.W. Sacks. *An Anthropologist on Mars: Seven Paradoxical Tales*. Vintage Books USA, 1996.

[62] G. Sapiro and D. Ringach. Anisotropic diffusion of multivalued images with applications to color filtering. *IEEE Trans. Image Processing*, 5(11):1582–1586, Oct 1996.

[63] C. Schmid and R. Mohr. Local grayvalue invariants for image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(5):530–534, May 1997.

[64] C. Schmid, R. Mohr, and C. Bauckhage. Evaluation of interest point detectors. *International Journal of Computer Vision*, 37(2):151–172, 2000.

[65] N. Sebe, Q. Tian, E. Loupias, M.S. Lew, and T.S. Huang. Evaluation of salient point techniques. *Image and Vision Computing*, 21(13-14):1087–1095, 2003.

[66] S.A. Shafer. Using color to seperate reflection components. *COLOR research and application*, 10(4):210–218, Winter 1985.

[67] J. Shi and C. Tomasi. Good features to track. In *IEEE conference on Computer Vision and Pattern Recognition*, 1994.

[68] E. P. Simoncelli, E.H. Adelson, and D.J. Heeger. Probability distributions of optical flow. In *IEEE conference on Computer Vision and Pattern Recognition*, pages 310–315, 1991.

[69] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval: the end of the early years. *IEEE trans. on pattern analysis and machine intelligence*, 22(12):1349–1380, 2000.

[70] S.M. Smith and J.M. Brady. Susan-a new approach to low level image processing. *International Journal of Computer Vision*, 23(1):45–78, 1997.

[71] H.M.G. Stokman and Th. Gevers. Density estimation for color images. *Journal of Electronic Imaging*, pages 17–23, 2001.

[72] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *Proc. of the Sixth International Conference on Computer Vision*, Bombay, India, January 1998.

[73] J. van de Weijer and Th. Gevers. Color mode filtering. In *Proc. Int. Conference on Image Processing*, Thessaloniki, Greece, Oct. 2001.

[74] J. van de Weijer, Th. Gevers, and J.M. Geusebroek. Color edge detection by photometric quasi-invariants. In *Int'l Conf. Computer Vision*, pages 1520–1526, Nice, France, 2003.

[75] J. van de Weijer, Th. Gevers, and J.M. Geusebroek. Edge and corner detection by photometric quasi-invariants. *IEEE Trans. Pattern Analysis and Machine Intelligence*, accepted 2004.

[76] J. van de Weijer, Th. Gevers, and A.W.M. Smeulders. Robust photometric invariant features from the color tensor. *IEEE Trans. Image Processing*, accepted 2004.

[77] J. van de Weijer and R. van den Boomgaard. Least squares and robust estimation of local image structure. *International Journal of Computer Vision*, accepted 2004.

[78] J. van de Weijer and R. van den Boomgaard. Local mode filtering. In *CVPR'01*, Kauai, Hawaii, USA, Dec. 2001.

[79] J. van de Weijer, L.J. van Vliet, P.W. Verbeek, and M. van Ginkel. Curvature estimation in oriented patterns using curvilinear models applied to gradient vector fields. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 23(9):1035–1042, 2001.

[80] R. van den Boomgaard. The Kuwahara-Nagao operator decomposed in terms of a linear smoothing and a morphological sharpening. In H. Talbot and R. Beare, editors, *Proceedings of the 6th International Symposium on Mathematical Morphology*, pages 283–292, Sydney, Australia, April 2002. CSIRO Publishing.

[81] R. van den Boomgaard and J. van de Weijer. On the equivalence of local-mode finding, robust estimation and mean-shift analysis as used in early vision tasks. In *ICPR 2002*, 2002.

[82] R. van den Boomgaard and J. van de Weijer. Robust estimation of orientation for texture analysis. In *The 2nd International workshop on texture analysis and synthesis*, pages 135—138, Copenhagen, 2002.

[83] M. van Ginkel, P.W. Verbeek, and L.J. van Vliet. Curvature estimation for overlapping curved patterns using orientation space. In *Proc. Fourth Conf. Advanced School for Computing and Imaging*, Lommel, Belgium, 1998.

[84] B. van Ginneken and B.M. ter Haar Romeny. Applications of locally orderless images. In *Proceedings Scale-Space 99*, pages 10–22, 1999.

[85] L. van Gool, T. Tuytelaars, and A. Turina. Local features for image retrieval. In *State-of-the-Art in Content-Based Image and Video Retrieval*, pages 21–41. Kluwer Academic Publishers, 2001.

[86] L.J. van Vliet and P.W. Verbeek. Curvature and bending energy in 2d and 3d digitized images. In *Proc. 8th Scandinavian Conf. Image Processing*, Tromso, Norway, 1993.

[87] L.J. van Vliet and P.W. Verbeek. Estimators for orientation and anisotropy in digitized images. In *Proc. First Conf. Advanced School for Computing and Imaging*, Heijen,Netherlands, 1995.

[88] L.J. van Vliet and P.W. Verbeek. Curvature estimation from oriented fields. In *Proc. 11th Scandinavian Conf. Image Processing*, Kangerlussuaq, Greenland, 1999.

[89] P.W. Verbeek. A class of sampling-error free measures in oversampled band-limited images. *Pattern Recognition Letters*, 3:287–292, 1985.

[90] P.W. Verbeek, L.J. van Vliet, and J. van de Weijer. Improved curvature and anisotropy estimation for curved line bundles. In *Int. Conf. on Pattern Recognition*, pages 528–533, Brisbane, Australia, 1998.

[91] J. Weicker. *Anisotropic Diffusion in Image Processing*. Tuebner Verlag, Stuttgart, 1997.

[92] J. Weickert and T. Brox. Diffusion and regularization of vector- and matrix valued images. In O. Scherzer M. Z. Nashed, editor, *Inverse Problems, Image Analysis and Medical Imaging*, volume 313, pages 252—268, 2002.

[93] L.R. Williams and K.K. Thornber. A comparision of measures for detecting natural shapes in cluttered backgrounds. *International Journal of Computer Vision*, 34(2/3):81–96, 1999.

[94] M. Worring and A.W.M. Smeulders. Digital curvature estimation. *CVGIP: Image Understanding*, 58:366–382, 1993.

Samenvatting

Vandaag de dag maakt het merendeel van de beeldberwerkings operaties slechts gebruik van de luminantie (de grijswaarden) en wordt de kleureninformatie onbenut gelaten, ondanks het feit dat een aanzienlijk deel van beelddata tegenwoordig in kleurenformaat is. Dit proefschrift behandelt technieken en theorieën om de mogelijkheden, die kleurenbeeldberwerking en in het bijzonder kleurenkenmerkdetectie bieden, verder te benutten.

Voor het gebruik van kleurenbeelden is het belangrijk om na te gaan hoe bestaande operaties, ontwikkeld voor luminantiebeelden, op een wiskundig correcte wijze naar kleurenbeelden kunnen worden uitgebreid. Voor beeldverfraaiing betekent dit dat de correlatie tussen de kanalen (rode, groene en blauwe kanaal) zodanig moet worden gerespecteerd dat de operaties geen kleuren introduceren die niet aanwezig zijn in het originele beeld. In dit proefschrift wordt dit probleem omgeschreven naar een robuust schattingsprobleem, en wordt een efficiënte methode voorgesteld om lokale beeldstructuur te schatten. Beeldverfraaiing gebaseerd op deze methode laat een goede ruisonderdrukking zien, gecombineerd met behoud van contrast en zonder de introductie van ongewenste nieuwe kleuren.

Voor operaties die gebaseerd zijn op de differentiële structuur van een beeld veroorzaakt de uitbreiding naar kleuren een ander wiskundig probleem: hoe moet de differentiële structuur van de verschillende kanalen worden gecombineerd ? Er wordt aangetoond dat tensor wiskunde dit probleem oplost, waarna een overzicht wordt gegeven van bestaande tensor gebaseerde operaties, samen met de uitbreiding voor kleurenbeelden.

Verder concentreert dit proefschrift zich op twee voordelen van kleurenbeeldbewerking ten opzichte van traditionele luminantie gebaseerde beeldbewerking.

Ten eerste, kleurenbeelden bevatten een rijkere fotometrische beschrijving van de beeldinhoud. Hierdoor wordt het mogelijk om fotometrisch invariante kenmerkdetectie in beelden te doen. Beeld operaties kunnen zo worden ontworpen dat belangrijke overgangen tussen objecten wel worden gedetecteerd terwijl relatief onbelangrijke schaduw randen worden genegeerd. Hoewel fotometrische invariantie algemeen gebruikt wordt, is er slechts weinig onderzoek gedaan naar uitbreiding van fotometrische invariantie theorie naar de differentiële structuur van beelden. In dit proefschrift stellen wij een groep afgeleidde filters voor, genaamd de quasi-invarianten, die het mogelijk maakt om de fotometrische invariante differentiaal structuur van beelden op een robuuste manier te meten. Experimenten tonen aan dat de quasi-invarianten

91

betere resultaten behalen dan de bestaande methodes wat betreft discriminerend vermogen en lokalisatie van de beeldkenmerken.

Een tweede voordeel van kleurenbeeldbewerking is dat kleur een belangrijke aanwijzing is voor saillante (in het oog springende) beeldpunten. Bestaande saillante beeldpuntoperaties zijn gebaseerd op de differentiële structuur van beelden, en maken geen gebruik van kleuren informatie. De saillantie van een beeldpunt wordt onder andere bepaald door zijn zeldzaamheid, omdat zeldzame beeldpunten meer informatie bevatten dan veel voorkomende beeldpunten. In dit proefschrift analyseren we de kansverdeling van kleurenafgeleiden voor een grote dataset van 40.000 beelden, en komen tot de observatie dat deze verdeling goed benaderd kan worden door een ellipsoïde. Deze opmerkelijke observatie wordt gebruikt om saillante beeldpuntdetectie te optimaliseren. Experimenten laten zien dat de hiervoor aangepaste methodes beeldpunten detecteren met hogere informatie dichtheid.

Dankwoord

Ik ben opgelucht dat ik de eindstreep heb gehaald, dit zou echter niet gelukt zijn zonder de hulp van velen. Ten eerste bedank ik mijn copromotor Theo Gevers voor zijn belangrijke bijdrage aan dit proefschrift. Vooral zijn positieve en op cruciale momenten aanmoedigende houding is van groot belang geweest bij de totstandkoming van dit proefschrift. Ook het regelmatige wakker schudden van mijn wat behoudende wetenschapsaanpak door Theo's onconformistische (zeg niet anarchistische) manier van wetenschap heeft een positieve invloed gehad op het eindresultaat.

Verder bedank ik mijn promotor Arnold Smeulders voor de stimulerende gesprekken, waar wetenschap werd afgewisseld met de kunst van het schrijven, en een kijkje in de universiteitspolitiek, waarbij ik vaak genoeg niet zeker wist op welke van die drie zijn laatste opmerking betrekking had. Verder, bedank ik Arnold voor zijn begrip en geduldige houding tijdens mijn lange periode van ziekte.

Rein van den Boomgaard bedank ik voor de inspirerende samenwerking die tot hoofdstuk 2 heeft geleid. Reins enorme vakkennis samen met zijn goed gestelde vragen leidde steevast tot nieuwe ideeën en diepere inzichten. Ook wil ik Jan-Mark Geusebroek bedanken voor onze samenwerking waarvan hoofdstuk 3 het bewijs is. Zijn enthousiasme en inzicht maakten het een genoegen om met hem samen te werken.

Verder bedank ik Piet Verbeek, Michael van Ginkel en in het bijzonder Lucas van Vliet, wier enthousiasme voor beeldbewerkingonderzoek op mij oversloeg. Hetgeen resulteerde in hoofdstuk 4 en mijn keuze voor deze tak van de wetenschap.

Van mijn collega's gaat speciale dank uit naar Andy die verantwoordelijk gehouden kan worden voor de resterende Engelse spelfouten (als niet terecht dan toch), en op wiens hulp van wiskundige tot LATEX problemen ik altijd heb kunnen rekenen. Veel dank. Ook volgde er op "Cuba ?", "Valentijn ?" of "East of Eden ?" altijd een "Sure.", en een tocht richting stad vaak in goed gezelschap van Frans, Aristeides, Cesar, Nikos, Frank, Jan, en Nicu. Verder bedank ik Frank (fj) voor zijn hulp, advies en het altijd klaarstaan voor een koffie-break. Dennis voor de programmeerhulp en de tennislessen. En bedank ik Gijs, Carlo, Harro, Cees, Gertjan, Sennay, Erik, Michiel, Marcel, Minh, Thang, Ioannis, Hieu, Giang en Frank (aldershó) voor hun hulp en sfeer in de groep.

Ten slotte dank ik mijn vrienden, mijn moeder, Jeroen en zusje voor hun steun, de nodige ontspanning en de interesse in mijn vooruitgang. Mochten zij denken dat ik na 6 jaar van "iets met beelden en kleur" deze fase van mijn leven wel heb afgerond, moet ik ze teleurstellen, ik heb nog zin in jaren van "iets met beelden".

# Color Plates

(a) (b) (c) (d) (e)
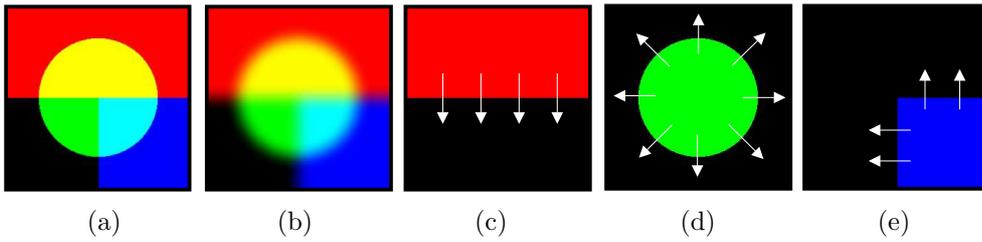
Figure C.1: *Chapter 1: (a) Example image and (b) linear smoothed version of example image. (c) Red channel, (d) green channel and (e) blue channel of example image.*


(a) (b) (c)
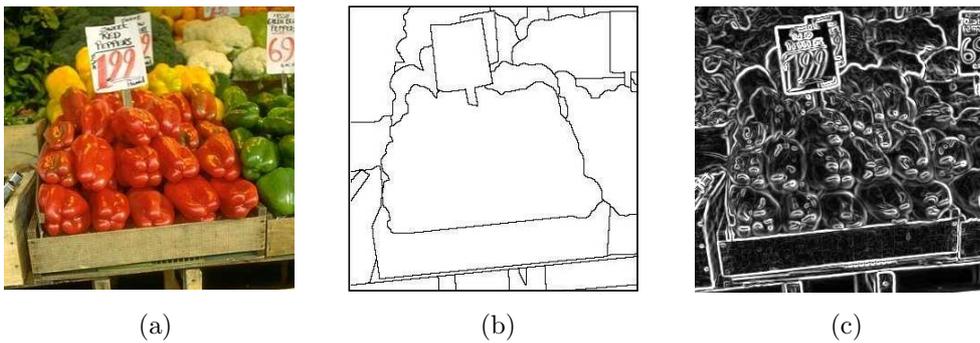
Figure C.2: *Chapter 1: (a) Example image, (b) human scene segmentation and (c) standard computer edge detection.*

$$(a) \qquad\qquad (b)$$
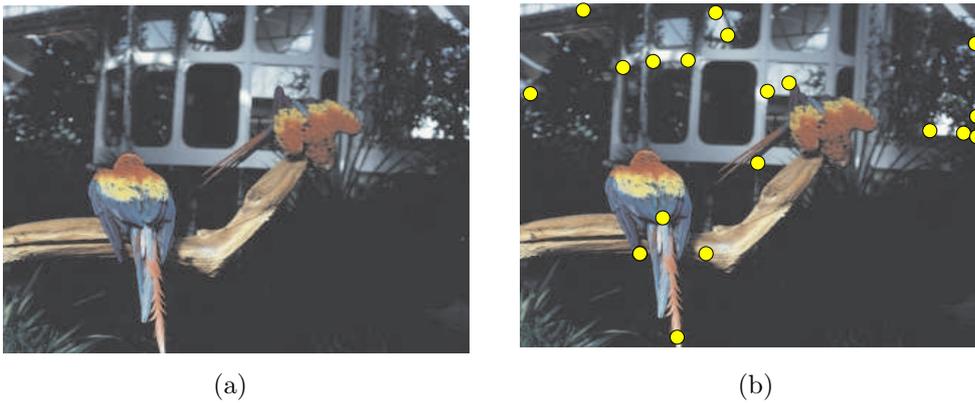
Figure C.3: *Chapter 1: (a) Example image, and (b) results of a standard salient point detector.*
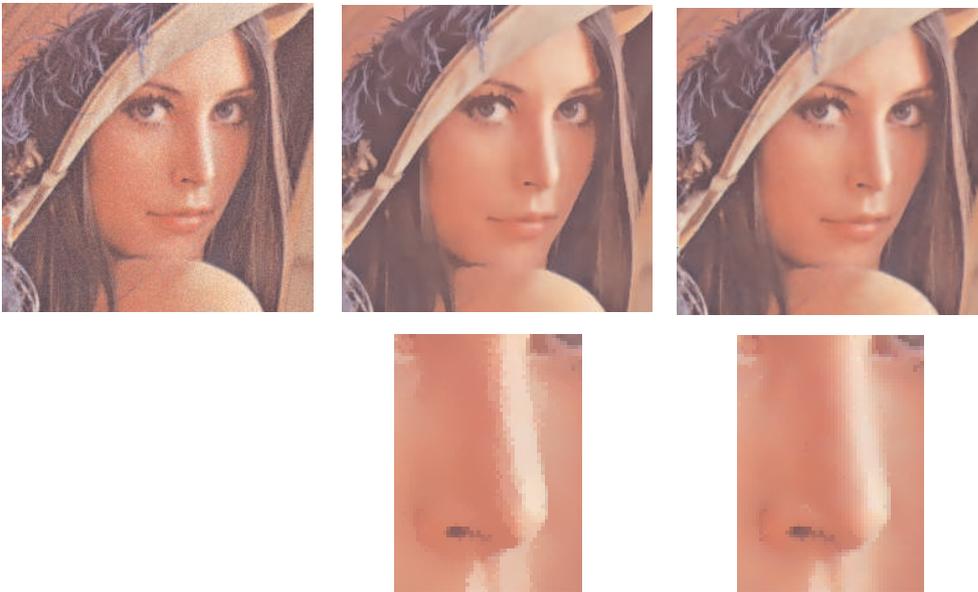
Figure C.4: *Chapter 2: Robust Estimation of Local Structure in Color Images. On the first row from left to right: the 'Lena' image with some noise added to it, the zero-order facet model based robust estimator of the values and the robust estimator based on a first order based facet model. On the second row we show a detail from the image above.*

(a)   (b)   (c)   (d)   (e)   (f)
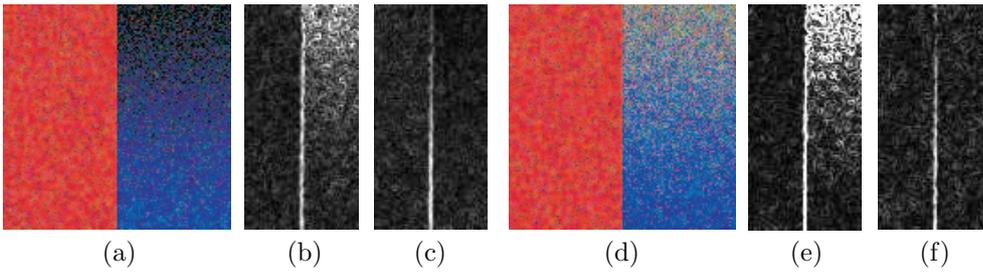
Figure C.5: *Chapter 3: (a) Red-blue edge, with a decreasing intensity of the blue patch going in the upward direction. Response of (b) normalized RGB derivative, and (c) shadow-shading quasi-invariant ($\mathbf{S}_x^c$). (d) Red-blue edge, with decreasing saturation going in the upward direction. Response of (e) hue derivative (h_x), and (f) specular-shadow-shading quasi-invariant ($\mathbf{H}_x^c$).*
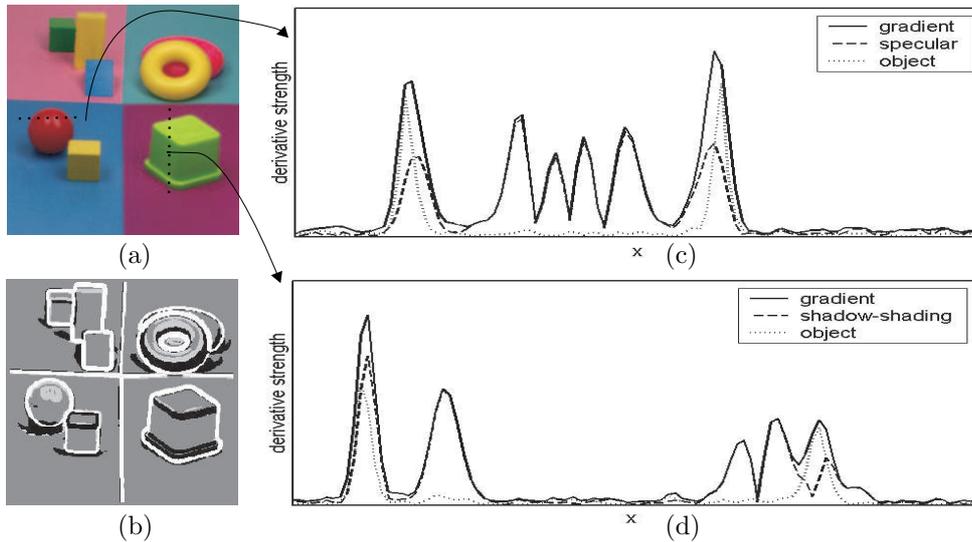
Figure C.6: *Chapter 3: (a) Input image with superimposed two dotted lines which are plotted in the images (c) and (d). (b) Edge classification result, with white object edges, black shadow edges and light grey specular edges. (c),(d) The derivative strength along lines indicated in (a).*
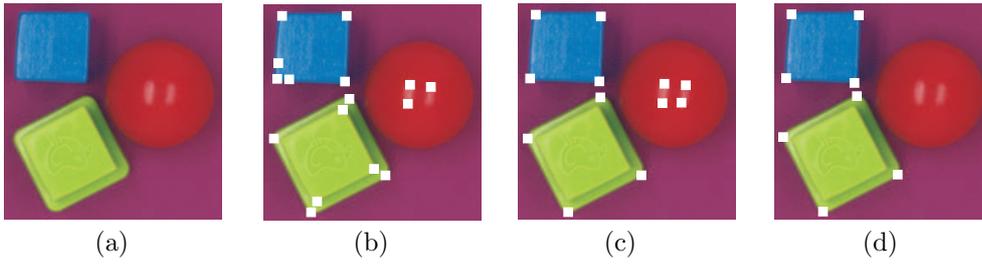
Figure C.7: *Chapter 3: (a) Input image and corner detector results based on (b) RGB gradient ($\mathbf{f}_x$), (c) shadow-shading quasi-invariant ($\mathbf{S}_x^c$), and (d) shadow-shading-specular quasi-invariant ($\mathbf{H}_x^c$).*
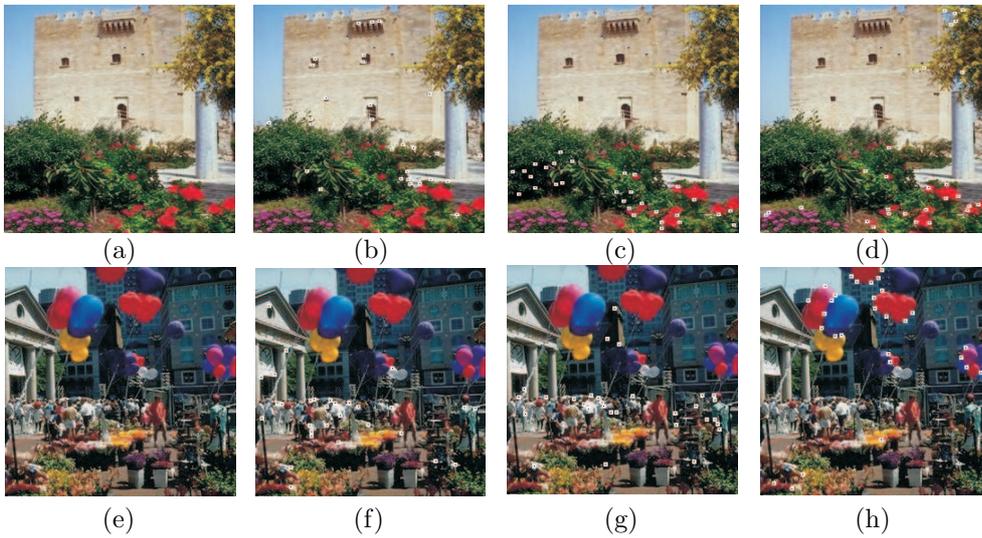
Figure C.8: *Chapter 3: (a), (e) Input images. Corner detection based on (b) RGB gradient ($\mathbf{f}_x$), (c) normalized RGB, (d) shadow-shading quasi-invariant ($\mathbf{S}_x^c$), (f) RGB gradient ($\mathbf{f}_x$), (g) hue full invariant (h_x), and (h) shadow-shading quasi-invariant ($\mathbf{H}_x^c$).*
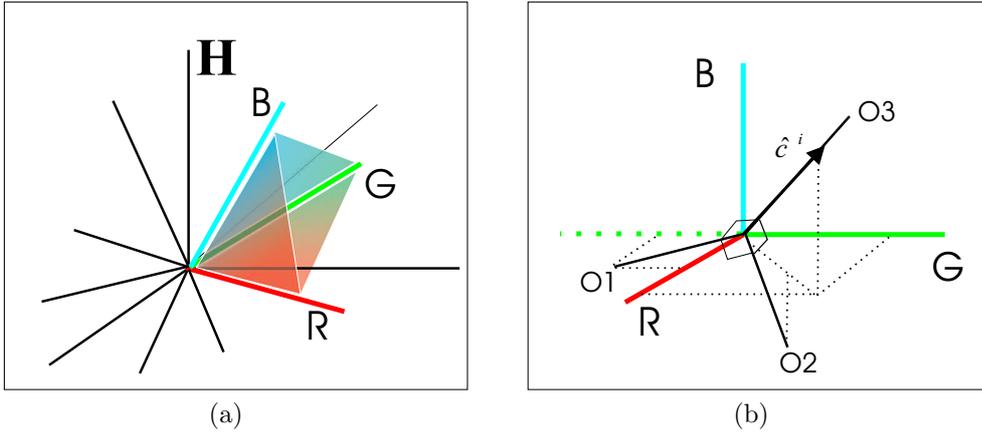
Figure C.9: *Chapter 5: (a) The subspace of measured light in the Hilbert space of possible spectra. (b) The RGB coordinate system and an alternative orthonormal color coordinate system which spans the same subspace.*
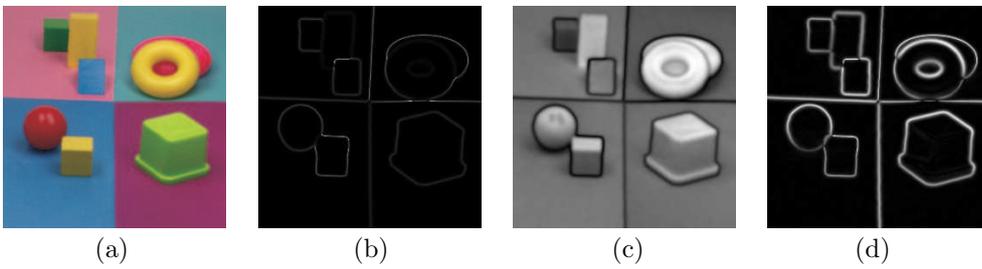
Figure C.10: *Chapter 5: (a) test image (b) hue derivative (c) saturation (d) quasi-invariant.*
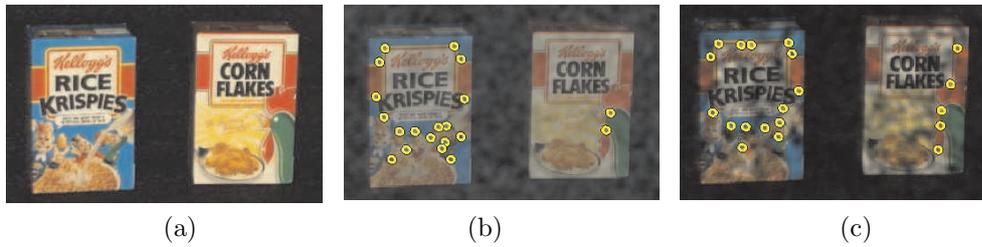
Figure C.11: *Chapter 5: (a) An example from Soil-47 image. (b) shadow-shading distortion with the shadow-shading quasi-invariant Harris points superimposed (c) specular distortion and the shadow-shading-specular Harris points superimposed.*
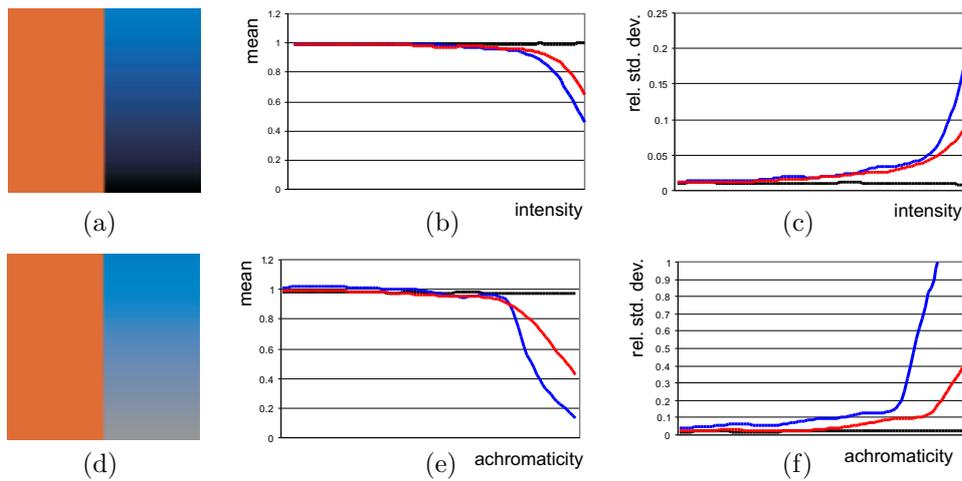
Figure C.12: *Chapter 5: (a),(d) frame from test sequence with constant optical flow of one pixel per frame. (b),(c) mean and relative standard deviation mean of the optical flow based on RGB (black line), shadow-shading invariant (blue line) and robust shadow-shading invariant (red line). (e),(f) mean and relative standard deviation of the optical flow based on RGB (black line), shadow-shading-specular invariant (blue line) and robust shadow-shading-specular invariant (red line).*
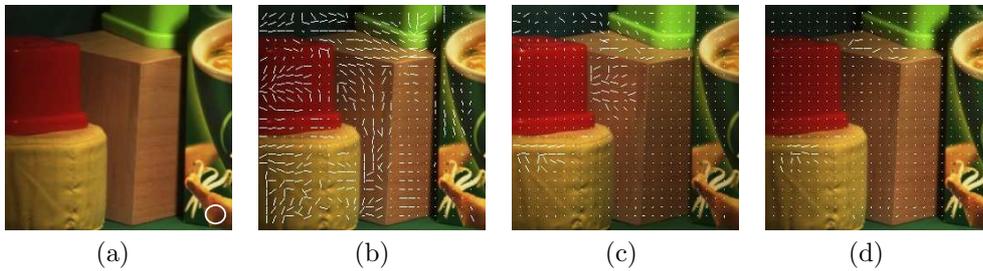
Figure C.13: *Chapter 5: (a) frame 1 of object scene with filter size superimposed on it. (b) RGB gradient optical flow (c) shadow-shading invariant optical flow and (d) robust shadow-shading invariant optical flow.*
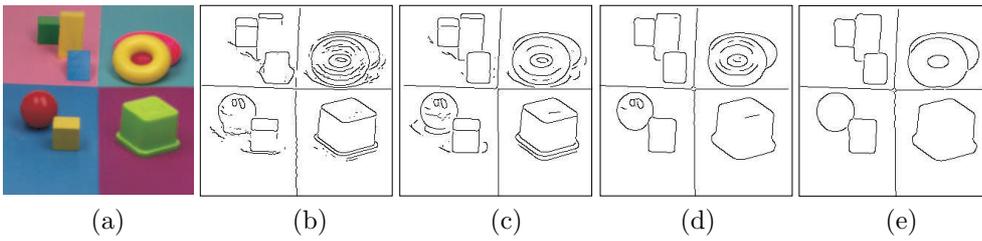
Figure C.14: *Chapter 5: (a) input image with Canny edge detection based on successively (b) luminance derivative (c) RGB derivatives (d) the shadow-shading quasi-invariant (e) the shadow-shading-specular quasi-invariant.*
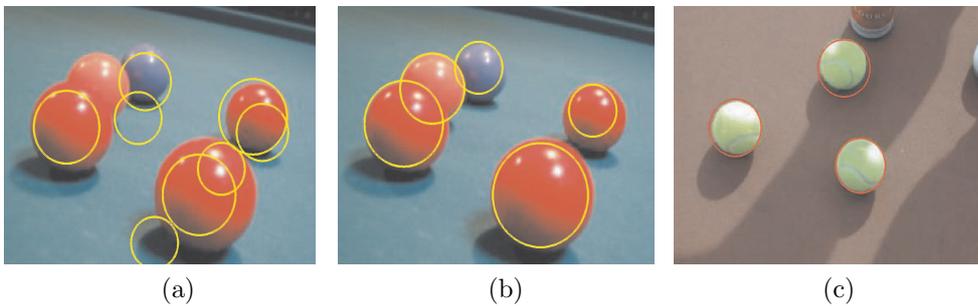

Figure C.15: *Chapter 5: (a) detected circles based on luminance (b) detected circles based on shadow-shading-specular quasi-invariant (c) detected circles based on shadow-shading-specular quasi-invariant.*
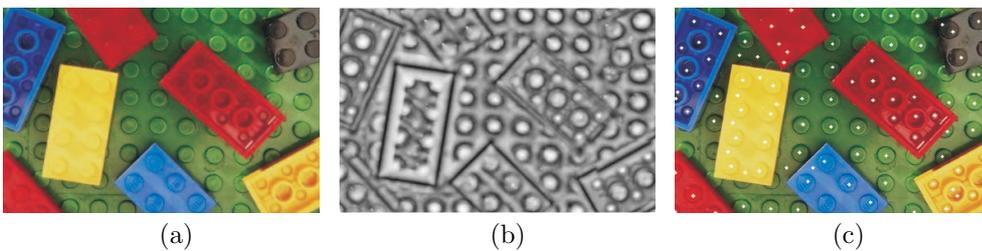

Figure C.16: *Chapter 5: (a) input image (b) the circularity coefficient C (c) the detected circles.*
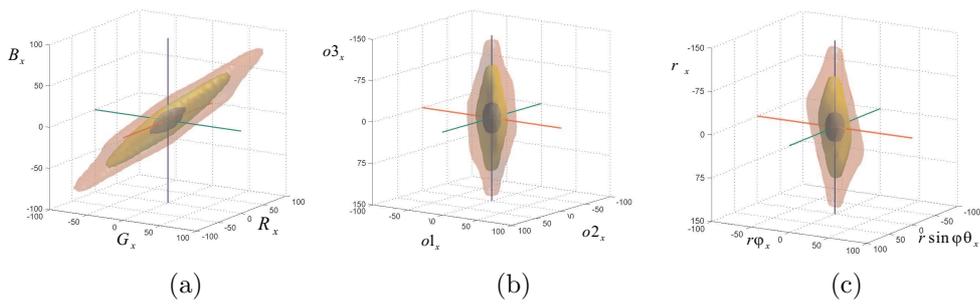
Figure C.17: *Chapter 6: The histograms of the distribution of the transformed derivatives of the Corel image database in respectively the (a) RGB coordinates, (b) the opponent coordinates and (c) the spherical coordinates. The three planes correspond with the isosalient surfaces which contain (from dark to light) respectively 90%, 99%, 99.9%t of the total number of pixels.*
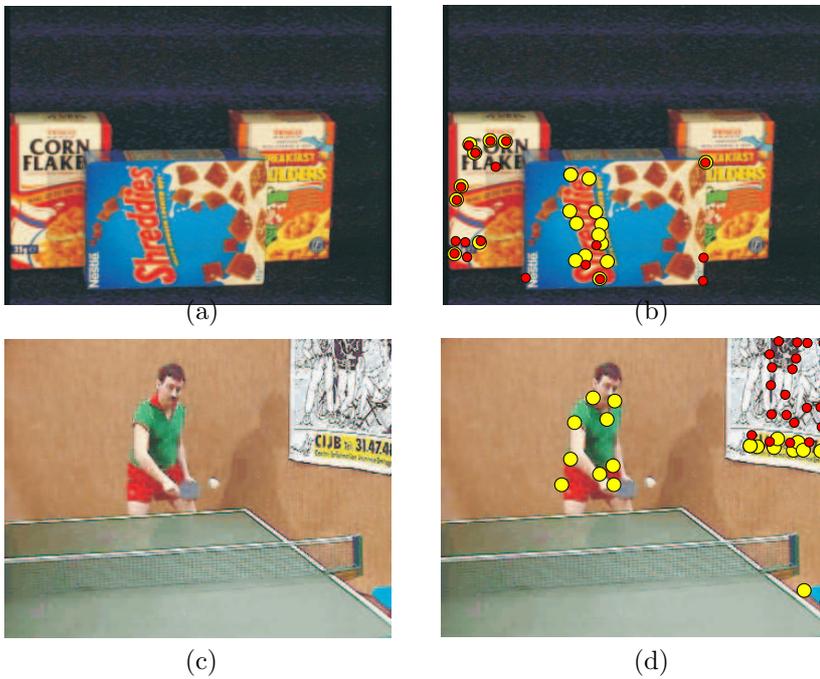

Figure C.18: *Chapter 6: (a) Example Soil data set and (c) frame from table-tennis sequence. (b) and (d) results of Harris detector (red dots) and the Harris detector with color boosting (yellow dots). The red dots mainly coincide with black and white events, while the yellow dots are focussed on colorful points.*

| (a) | (b) | (c) | (d) |
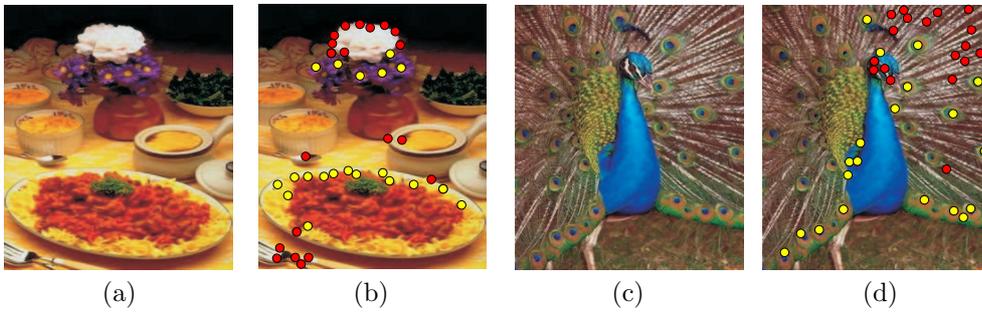
Figure C.19: *Chapter 6: (a) and (c) Corel input images. (b) and (d) results of Harris detector (red dots) and the Harris detector with color boosting (yellow dots). The red dots mainly coincide with black and white events, while the yellow dots are focussed on colorful points.*


| (a) | (b) | (c) |
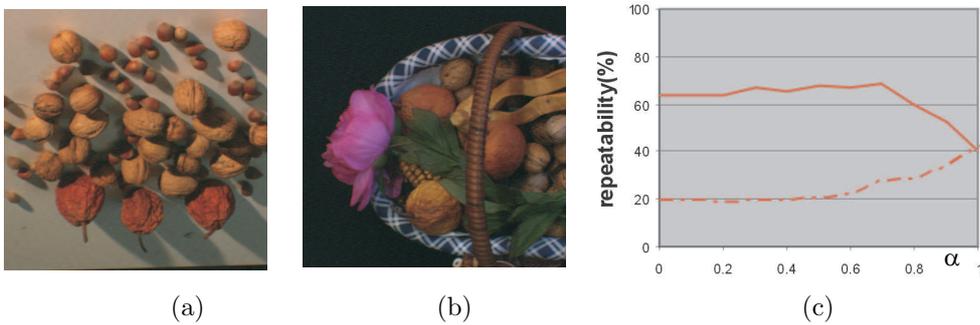
Figure C.20: *Chapter 6: (a),(b) Two frames from two sequences with changing illumination conditions. (c) Repeatability as a function of the amount of color saliency boosting for the two sequences. Dotted line for the nuts-sequence and the continuous line for the fruit-basket sequence.*
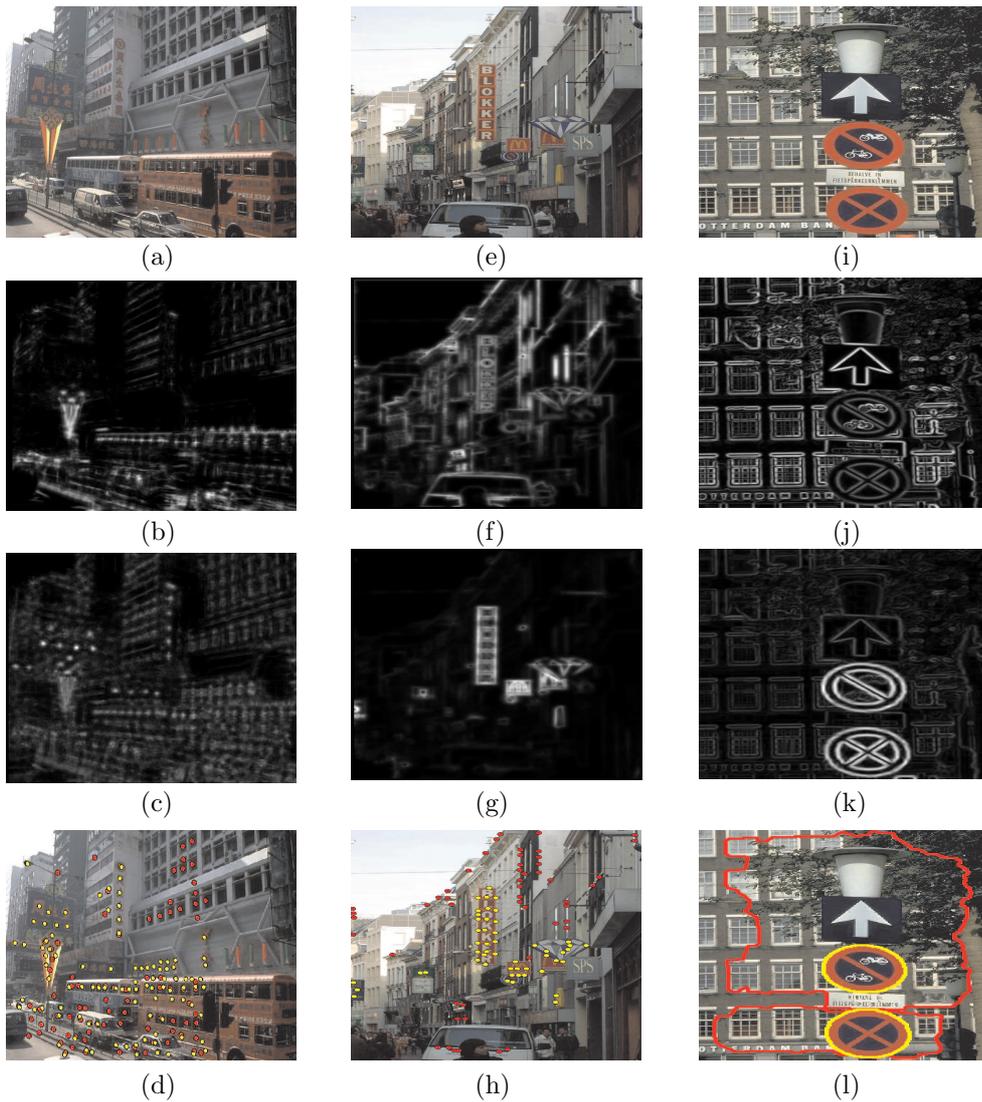
Figure C.21: *Chapter 6: Respectively, the input image, RGB-gradient based saliency map, the color boosted saliency map and the results with red dots (lines) for the gradient-based method and yellow dots (lines) for the salient points after color saliency boosting. (a),(b),(c),(d) Results for the focus points, (e),(f),(g),(h) for the symmetry points and (i),(j),(k),(l) for the global optimal regions and boundary method.*