



# 1 Introduction

Color is one of the most important and fascinating aspects of the world surrounding us. To comprehend the broad characteristics of color, a range of research fields has been actively involved, including physics (light and reflectance modeling), biology (visual system), physiology (perception), linguistics (cultural meaning of color), and art.

From a historical perspective, covering more than 400 years, prominent researchers contributed to our present understanding of light and color. Snell and Descartes (1620–1630) formulated the law of light refraction. Newton (1666) discovered various theories on light spectrum, colors, and optics. The perception of color and the influence on humans has been studied by Goethe in his famous book “*Farbenlehre*” (1840). Young and Helmholtz (1850) proposed the trichromatic theory of color vision. Work on light and color resulted in quantum mechanics elaborated by Max Planck, Albert Einstein, and Niels Bohr. In art (industrial design), Albert Munsell (1905) invented the theory on color ordering in his “*A Color Notation*.” Further, the value of the biological and therapeutic effects of light and color have been analyzed, and views on color from folklore, philosophy, and language have been articulated by Schopenhauer, Hegel, and Wittgenstein.

Over the last decades, with the technological advances of printers, displays, and digital cameras, an explosive growth in the diversity of needs in the field of color computer vision has been witnessed. More and more, the traditional gray value imaginary is replaced by color systems. Moreover, today, with the growth and popularity of the World Wide Web, a tremendous amount of visual information, such as images and videos, has become available. Hence, nowadays, all visual data is available in color. Furthermore, (automatic) image understanding is becoming indispensable to handle large amount of visual data. Computer vision deals with image understanding and search technology for the management of

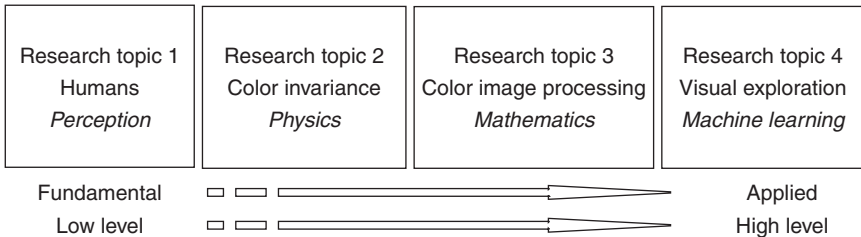
large-scale pictorial datasets. However, in computer vision, the use of color has been only partly explored so far.

This book determines the use of color in computer vision. We take on the challenge of providing a substantial set of color theories, computational methods, and representations, as well as data structures for image understanding in the field of computer vision. Invariant and color constant feature sets are presented. Computational methods are given for image analysis, segmentation, and object recognition. The feature sets are analyzed with respect to their robustness to noise (e.g., camera noise, occlusion, fragmentation, and color trustworthiness), expressiveness, discriminative power, and compactness (efficiency) to allow for fast visual understanding. The focus is on deriving semantically rich color indices for image understanding. Theoretical models are presented to express semantics from both a physical and a perceptual point of view.

## 1.1 From Fundamental to Applied

---

The aim of this book is to present color theories and techniques for image understanding from (low level) basic color image formation to (intermediate level) color invariant feature extraction and color image processing to (high level) learning of object and scene recognition by semantic detectors. The topics, and corresponding chapters, are organized from low level to high level processing and from fundamental to more applied research. Moreover, each topic is driven by a different research area using color as an important stand-alone research topic and as a valuable collaborative source of information bridging the gap between different research fields (Fig. 1.1).



**Figure 1.1** The different topics are organized from low level to high level processing and from fundamental to more applied research. Each topic is driven by a different research area from human perception, physics, and mathematics to machine learning.

The book starts with the explanation of the mechanisms of *human color perception*. Understanding the human visual pathway is crucial for computer vision systems, which aim to describe color information in such a way that it is relevant to humans.

Then, *physical* aspects of color are studied, resulting in reflection models from which photometric invariance is derived. Photometric invariance is important for computer vision, as it results in color measurements that are independent of accidental imaging conditions such as a change in camera viewpoint or a variation in the illumination.

A *mathematical* perspective is taken to cope with the difference between gray value (scalar) and color (vector) information processing, that is, the extension of single-channel signal to multichannel signal processing. This mathematical approach will result in a sound way to perform color processing to obtain (low level) computational methods for (local) feature computation (e.g., color derivatives), descriptors (e.g., SIFT), and image segmentation. Furthermore, based on both mathematical and physical fundamentals, color image feature extraction is presented by integrating differential operators and color invariance.

Finally, color is studied in the context of *machine learning*. Important topics are color constancy, photometric invariance by learning, and color naming in the context of object recognition and video retrieval. On the basis of the multichannel approach and color invariants, computational methods are presented to extract salient image patches. From these salient image patches, color descriptors are computed. These descriptors are used as input for various machine learning methods for object recognition and image classification.

The book consists of five parts, which are discussed next.

## 1.2 Part I: Color Fundamentals

---

The observed color of an object depends on a complex set of imaging conditions. Because of the similarity in trichromatic color processing between humans and computer vision systems, in Chapter 2, an outline on human color vision is provided. The different stages of color information processing along the human visual pathway are presented. Further, important chromatic properties of the visual system are discussed such as chromatic adaptation and color constancy. Then, to provide insights in the imaging process, in Chapter 3, the basics on color image formation are presented. Reflection models are introduced describing the imaging process and how photometric changes, such as shadows and specularities, influence the RGB values in an image. Additionally, a set of relevant color spaces are enumerated.

## 1.3 Part II: Photometric Invariance

---

In computer vision, invariant descriptions for image understanding are relatively new but quickly gaining ground. The aim of photometric invariant features is to compute image properties of objects irrespective of their recording conditions.

This comes, in general, at the loss of some discriminative power. To arrive at invariant features, the imaging process should be taken into account.

In Chapters 4–6, the aim is to extract color invariant information derived from the physical nature of objects in color images using reflection models. Reflection models are presented to model dull and gloss materials, as well as shadows, shading, and specularities. In this way, object characteristics can be derived (based on color/texture statistics) for the purpose of image understanding. Physical aspects are investigated to model and analyze object characteristics (color and texture) under different viewing and illumination conditions. The degree of invariance should be tailored to the recording circumstances. In general, a color model with a very wide class of invariance loses the power to discriminate among object differences. Therefore, in Chapter 6, the aim is to select the tightest set of invariants suited for the expected set of nonconstant conditions.

### 1.3.1 Invariance Based on Physical Properties

As discussed in Chapter 4, most of the methods to derive photometric invariance are using 0th order photometric information, that is, pixel values. The effect of the reflection models on higher-order- or differential-based algorithms remained unexplored for a long time. The drawbacks of the photometric invariant theory (i.e., the loss of discriminative power and deterioration of noise characteristics) are inherited by the differential operations. To improve the performance of differential-based algorithms, the stability of photometric invariants can be increased through the noise propagation analysis of the invariants. In Chapters 5 and 6, an overview is given on how to advance the computation of both photometric invariance and differential information in a principled way.

### 1.3.2 Invariance By Machine Learning

While physical-based reflection models are valid for many different materials, it is often difficult to model the reflection of complex materials (e.g., with nonperfect Lambertian or dielectrical surfaces) such as human skin, cars, and road decks. Therefore, in Chapter 7, we also present techniques to estimate photometric invariance by machine learning models. On the basis of these models, computational methods are studied to derive the (in)sensitivity of transformed color channels to photometric effects obtained from a set of training samples.

## 1.4 Part III: Color Constancy

---

Differences in illumination cause measurements of object colors to be biased toward the color of the light source. Humans have the ability of color constancy; they tend to perceive stable object colors despite large differences in illumination. A similar color constancy capability is necessary for various computer

vision applications such as image segmentation, object recognition, and scene classification.

In Chapters 8–10, an overview is given on computational color constancy. Many state-of-the-art methods are tested on different (freely) available datasets. As color constancy is an underconstrained problem, color constancy algorithms are based on specific imaging assumptions. These assumptions include the set of possible light sources, the spatial and spectral characteristics of scenes, or other assumptions (e.g., the presence of a white patch in the image or that the averaged color is gray). As a consequence, no algorithm can be considered as universal. With the large variety of available methods, the inevitable question, that is, how to select the method that induces the equivalence class for a certain imaging setting, arises. Furthermore, the subsequent question is how to combine the different algorithms in a proper way. In Chapter 11, the problem of how to select and combine different methods is addressed. An evaluation of color constancy methods is given in Chapter 12.

## 1.5 Part IV: Color Feature Extraction

---

We present how to extend luminance-based algorithms to the color domain. One requirement is that image processing 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. Therefore, the implications on the multichannel theory are investigated, and algorithmic extensions for luminance-based feature detectors such as edge, curvature, and circular detectors are given. Finally, the photometric invariance theory described in earlier parts of the book is applied to feature extraction.

### 1.5.1 From Luminance to Color

The aim is to take an algebraic (vector based) approach to extend scalar-signal to vector-signal processing. However, a vector-based approach is accompanied by several mathematical obstacles. Simply applying existing luminance-based operators on the separate color channels, and subsequently combining them, will fail because of undesired artifacts.

As a solution to the opposing vector problem, for the computation of the color gradient, the color tensor (structure tensor) is presented. In Chapter 13, we give a review on color-tensor-based techniques on how to combine derivatives to compute local structures in color images in a principled way. Adaptations of the tensor lead to a variety of local image features, such as circle detectors, curvature estimation, and optical flow.

## 1.5.2 Features, Descriptors, and Saliency

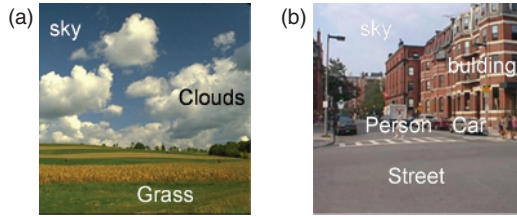
Although color is important to express saliency, the explicit incorporation of color distinctiveness into the design of image feature detectors has been largely ignored. To this end, we give an overview on how color distinctiveness can be explicitly incorporated in the design of color (invariant) representations and feature detectors. The approach is based on the analysis of the statistics of color derivatives. Furthermore, we present color descriptors for the purpose of object recognition. Object recognition aims to detect high level semantic information present in images and videos. The approach is based on salient visual features and using machine learning to build concept detectors from annotated examples. The choice of features and machine learning algorithms is of great influence on the accuracy of the concept detector. Features based on interest regions, also known as *local features*, consist of an interest region detector and a region descriptor. In contrast to the use of intensity information only, we will present both interest point detection (Chapter 13) and region description (Chapter 14), see Figure 1.2.



**Figure 1.2** Visual exploration is based on the paradigm to divide the images into meaningful parts from which features are computed. Saliient point detection is applied first from which color descriptors are computed. Then, machine learning is applied to provide classifiers for object recognition.

## 1.5.3 Segmentation

In computer vision, texture is considered as all what is left after color and local shape have been considered or it is given in terms of structure and randomness. Many common textures are composed of small textons usually too large in number to be perceived as isolated objects. In Chapter 15, we give an overview on powerful features based on natural image statistics or general principles from surface physics in order to classify a large number of materials by their texture. On the basis of their textural nature, different materials and concepts containing certain types of material can be identified (Fig. 1.3). For features at the level of (entire) objects, the aim is to aggregate pieces of local visual information to characteristic geographical arrangements of (possibly missing)



**Figure 1.3** On the basis of their textural nature, different materials and concepts containing certain types of material can be identified.

parts. The objective is to find computational models to combine individual observations of an object's appearance under the large number of variations in that appearance.

## 1.6 Part V: Applications

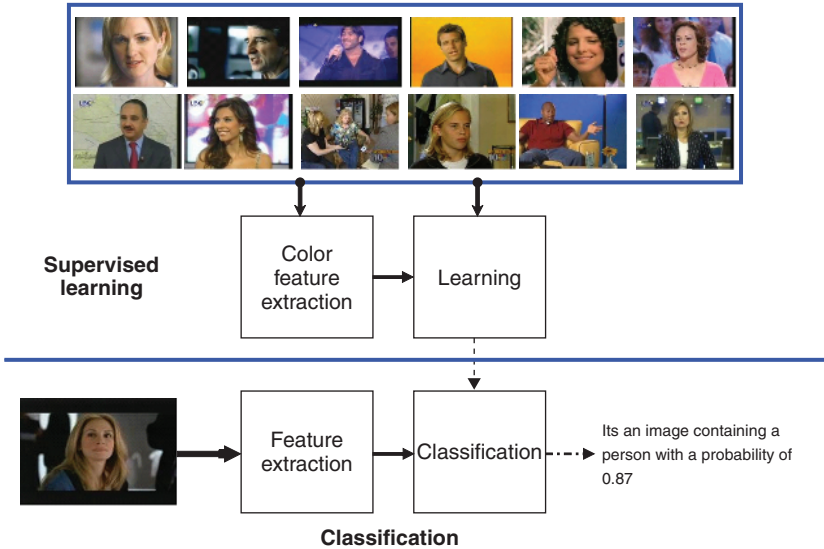
In the final part of the book, we emphasize on the importance of color in several computer vision applications.

### 1.6.1 Retrieval and Visual Exploration

In Chapter 16, we follow the state-of-the-art object recognition paradigm consisting of a learning phase and a (runtime) classification phase (Fig. 1.4). The learning module consists of color feature extraction and supervised learning strategies. Color descriptors are computed at salient points in the image by different point detectors (Fig. 1.2). The learning part is executed offline. The runtime classification part takes an image or video as an input from which features are extracted. Then, the classification scheme will provide a probability to what class of concepts the query image/video belongs to (people, mountain, or cars). A concept is defined as a *material* (e.g., grass, brick, or sand, as illustrated in Fig. 1.3a) or as an object (e.g., car, bike, or person, as illustrated in Fig. 1.3b), an event (explosion, crash, etc.), or a scene (e.g., mountain, beach, or city), see Figure 1.5.

### 1.6.2 Color Naming

Color names are linguistic labels that humans attach to colors. We use them routinely and seemingly without effort to describe the world around us. They have been primarily studied in the fields of visual psychology, anthropology, and linguistics. One of the most influential works in color naming is the linguistic study of Berlin and Kay on basic color terms. In Chapter 17, color names are



**Figure 1.4** First, during training, features are extracted and objects/scenes are learned offline by giving examples of different concepts (e.g., people, buildings, mountains) as the input to a learning system (in this case pictures containing people). Then, during online recognition, features are extracted from the incoming image/video and provided to the classification system to result in a probability of being one of the concepts.



**Figure 1.5** TRECVID concepts and corresponding key frames.

presented in the context of image retrieval. This allows for searching objects in images by a certain color name.

### 1.6.3 Multispectral Applications

Finally, in Chapter 18, we give an overview on multispectral imagery and applications to segmentation and detection. In fact, techniques are presented to detect regions in multispectral images. To obtain robustness against noise, noise propagation is adopted.



## 1.7 Summary

---

Visual information (images and video) is one of the most valuable sources of information. In fact, it is the core of current technologies such as the Internet and mobile phones. The immense stimulus of the use and exploitation of digital visual information demands for advanced knowledge representations, learning systems, and image understanding techniques. As all digital information is nowadays available in color (documents, images, videos, and movies), there is an increasing demand for the use and understanding of color information.

Although color has been proved to be a central topic in various disciplines, it has only been partly explored so far in computer vision, which this book resolves. The central topic of this book is to present color theories, color representation models, and computational methods, which are essential for visual understanding in the field of computer vision. Color is taken as the merging topic between different research areas such as mathematics, physics, machine learning, and human perception. Theoretical models are studied to express color semantics from both a physical and a perceptual point of view. These models are the foundations for visual exploration, which are tested in practice.

