UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 876

Spectral Image Processing with Applications in Biotechnology and Pathology

MILAN GAVRILOVIC

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2011

ISSN 1651-6214
ISBN 978-91-554-8209-1
urn:nbn:se:uu:diva-160574

Dissertation presented at Uppsala University to be publicly examined in Polhemsalen, Ångströmlaboratoriet, Lägerhyddsvägen 1, Uppsala. Friday, December 2, 2011 at 13:15 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract
Gavrilovic, M. 2011. Spectral Image Processing with Applications in Biotechnology and Pathology. Acta Universitatis Upsaliensis. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 876. 63 pp. Uppsala. ISBN 978-91-554-8209-1.

Color theory was first formalized in the seventeenth century by Isaac Newton just a couple of decades after the first microscope was built. But it was not until the twentieth century that technological advances led to the integration of color theory, optical spectroscopy and light microscopy through spectral image processing. However, while the focus of image processing often concerns modeling of how images are perceived by humans, the goal of image processing in natural sciences and medicine is the objective analysis. This thesis is focused on color theory that promotes quantitative analysis rather than modeling how images are perceived by humans.

Color and fluorescent dyes are routinely added to biological specimens visualizing features of interest. By applying spectral image processing to histopathology, subjectivity in diagnosis can be minimized, leading to a more objective basis for a course of treatment planning. Also, mathematical models for spectral image processing can be used in biotechnology research increasing accuracy and throughput, and decreasing bias.

This thesis presents a model for spectral image formation that applies to both fluorescence and transmission light microscopy. The inverse model provides estimates of the relative concentration of each individual component in the observed mixture of dyes. Parameter estimation for the model is based on decoupling light intensity and spectral information. This novel spectral decomposition method consists of three steps: (1) photon and semiconductor noise modeling providing smoothing parameters, (2) image data transformation to a chromaticity plane removing intensity variation while maintaining chromaticity differences, and (3) a piecewise linear decomposition combining advantages of spectral angle mapping and linear decomposition yielding relative dye concentrations.

The methods described herein were used for evaluation of molecular biology techniques as well as for quantification and interpretation of image-based measurements. Examples of successful applications comprise quantification of colocalization, autofluorescence removal, classification of multicolor rolling circle products, and color decomposition of histological images.

Keywords: color theory, light microscopy, spectral imaging, image analysis, digital image processing, mathematical modeling, estimation, noise models, spectral decomposition, color decomposition, colocalization, cross-talk, autofluorescence, tissue separation, prostate cancer, biomedical applications, molecular biotechnology, histopathology

Milan Gavrilovic, Uppsala University, Department of Information Technology, Centre for Image Analysis, Lägerhyddsv. 2, SE-752 37 Uppsala, Sweden.

ISSN 1651-6214
ISBN 978-91-554-8209-1
urn:nbn:se:uu:diva-160574 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-160574)

To my dear parents and my lovely wife

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I **M. Gavrilovic**, C. Wählby (2009) Quantification of colocalization and cross-talk based on spectral angles. *Journal of Microscopy*, 234(3):321-324

II **M. Gavrilovic**, C. Wählby (2009) Suppression of autofluorescence based on fuzzy classification by spectral angles. Proceedings of workshop associated with MICCAI: *Optical Tissue Image Analysis in Microscopy, Histopathology and Endoscopy*, pp.135-146

III **M. Gavrilovic**, I. Weibrecht, T. Conze, O. Söderberg, C. Wählby (2011) Automated classification of multicolored rolling circle products in dual-channel wide-field fluorescence microscopy. *Cytometry Part A*, 79A(7):518-527
Presented methods were used for evaluation of molecular biology techniques in the following paper:
I. Weibrecht, **M. Gavrilovic**, L. Lindbom, U. Landegren, C. Wählby, O. Söderberg (2011) Visualising individual sequence-specific protein-DNA interactions *in situ*. To appear in *New Biotechnology*

IV **M. Gavrilovic**, J. Azar, J. Lindblad, C. Wählby, E. Bengtsson, C. Busch, I. Carlbom (2011) Blind color decomposition of histological images. *manuscript for journal publication*

The author designed models and methods, and was the principal author in Papers I-IV. Irene Weibrecht and Tim Conze prepared biological specimens for Paper III. Christer Busch designed staining protocols for Paper IV.

Related work by the author

In the process of performing the research leading to this thesis, the author has also contributed to the following publications.

Patent applications

E. Bengtsson, **M. Gavrilovic**, J. Lindblad, C. Wählby (2008) *Pixel classification in image analysis*, US patent application 2009/214114

J. Azar, C. Busch, I. Carlbom, **M. Gavrilovic** (2011) *Color decomposition in histopathology*, US patent application (submitted)

The inventors are listed in alphabetical order.

Conference papers and abstracts

M. Gavrilovic, C. Wählby. Quantification and localization of colocalization, *Proceedings of Swedish Symposium in Image Analysis (SSBA) 2007*, pp.93-96

M. Gavrilovic, C. Wählby, J. Lindblad, E. Bengtsson. Algorithms for cross-talk suppression in fluorescence microscopy, *Abstracts of Medicinteknikdagarna 2008*, pp.64

M. Gavrilovic, C. Wählby, J. Lindblad, E. Bengtsson. Dimensionality Reduction for Colour Based Pixel Classification, *Proceedings of Swedish Symposium in Image Analysis (SSBA) 2009*, pp.65-68

M. Gavrilovic, C. Wählby, J. Lindblad, E. Bengtsson. Spectral Angle Histogram - a Novel Image Analysis Tool for Quantification of Colocalization and Cross-talk, *Proceedings of the 9th European Light Microscopy Initiative (ELMI) meeting 2009*, pp.66-67

M. Gavrilovic, J. Azar, C. Busch, I. Carlbom. Tissue Separation for Quantitative Malignancy Grading of Prostate Cancer, *Abstracts of Medicinteknikdagarna 2011*, pp.32

Contents

Notation

x	italics denote scalars and scalar-valued functions
$\mathbf{x}$	boldface lower-case letters denote vectors
$\mathbf{X}$	boldface upper-case letters denote matrices
$\mathbb{R}$	set of real numbers
$\mathbb{Z}$	set of integer numbers
ι	imaginary unit
$[\cdot]^T$	vector or matrix transpose
$\|\cdot\|_p$	p-norm of the vector
$\ln(\cdot)$	the natural logarithm
$\exp(\cdot)$	the exponential function
	Logarithm and exponential functions of a vector are applied element-wise.
$\mathbf{p}$	imaged position in $\mathbb{R}^3$ space, $\mathbf{p} \in \mathbb{Z}^3$
$n_\mathbf{p}$	number of imaged positions
λ	wavelength
$c_\mathbf{p}(\lambda)$	spectral signature at position $\mathbf{p}$
n_c	number of spectral channels
$k = 1, \ldots, n_c$	index over the n_c spectral channels
$\mathbf{c_p}$	sampled spectral signature at position $\mathbf{p}$
$\mathbf{s_p}$	spectral image
n	number of dyes
$j = 1, \ldots, n$	index over the n dyes
$\mathbf{d_p}$	relative dye concentrations at position $\mathbf{p}$
$\mathbf{a}_j$	estimated sampled spectral signature of a dye j
$\mathbf{A}$	mixing matrix

1. Introduction

Motivation

With the twentieth century technological advancements, otherwise closely related fields of color theory, spectroscopy, microscopy and image processing were merged in a true synergy. In parallel, a growing understanding of the human visual information processing created space for a quantitative presentation of color as compared to that perceived by humans. Having in mind that the aim of image processing in natural sciences and medicine is to secure objective analysis, this thesis relies on fundamentals of color theory that promote quantitative analysis instead of the subjective approach that prevails in the field today. Guided by this, the purpose of incorporating color and spectral image processing into histopathology is to minimize subjectiveness and increase the reliability of diagnostics. Further on, image processing benefits from of mathematical models that can, with maximum reliability, confirm or disapprove scientific hypotheses at an early stage in research.

The aim of this thesis is to present a unified framework for processing of microscopy images based on decoupling light intensity and spectral information. The method comprises a mathematical model and algorithms for automated identification of its parameters. It deals with a number of light microscopy applications important for quantitative analysis:

- Suppression of cross-talk (bleed-through) in fluorescence microscopy
- Suppression of background fluorescence (autofluorescence)
- Detection of colocalization
- Color decomposition of histological images in bright-field microscopy, both well-conditioned and ill-conditioned cases.

Thesis outline

The central motive of the thesis and the four included papers is quantitative analysis of color and spectral images with applications in microscopy.

Chapter 2 presents the theoretical fundamentals of the problem from different perspectives: What is an image? What is color? How to acquire spectral images? How image data is processed and what scientists in related fields can do to facilitate the process?

Chapter 3 derives a linear model from the Beer-Lambert law and describes how related spectral image processing methods estimate its param-

eters. Chapter 4 introduces novel chromaticity spaces and describes how they can be employed to solve two main problems: estimation of the model parameters and evaluation of the dyes. The chapter ends with a description of a piecewise linear decomposition algorithm designed to solve ill-conditioned problems.

Chapter 5 describes several applications where the method was successfully applied. The chapter also includes brief summaries of the included papers and the most important conclusions. Finally, Chapter 6 is dedicated to describing limitations of the model and the method and discussion concerning the future development of the field.

2. Background

Alas, I have studied philosophy,
the law as well as medicine,
and to my sorrow, theology;
studied them well with ardent zeal,
yet here I am, a wretched fool,
no wiser than I was before.
– Faust,
from the homonymous play by
Johann Wolfgang von Goethe (1749-1832)

2.1 Digital images

Image processing and related fields

In electrical engineering, sensors are devices that measure and convert physical quantities to analogue signals. Analogue signals contain continuous spatial, time-varying, or spectral information and have continuous magnitude. In contemporary sensors, electric circuits sample analogue signals and analogue-to-digital converters quantize those magnitudes, hence forming digital signals represented as finite sequences. Signal processing comprises a group of techniques that aim to represent and transform these finite sets of input measurements from sensors (i.e., signals) using some *useful* operation. For instance, a *useful* signal processing operation could increase the signal-to-noise ratio.

Image processing is a field concerned with transforming the input image. Similarly to the definition of signals, here the image is a discrete representation of the quantized spatial energy distribution of a source of radiant energy [1]. If the sensor by any means succeeds in measuring physical quantities of interests at spatial positions $\mathbf{p} = (p_x, p_y, p_z)$, it is common to define an image as a discrete function over three-dimensional discrete space, giving discrete geometry a fundamentally intrinsic role in image processing. For example, it is not trivial to define a digital straight line [2].

Historically, early image processing techniques followed the invention of television, but the field started to develop at a higher pace in the 1960's with sufficiently powerful computers, advancements in satellite imagery, medical imaging and the invention of CCD cameras. As humans visualize

and comprehend images with less difficulty than raw numerical data from sensors, a new field was born - image analysis. The aim of image analysis techniques is to extract *useful* information from the input image, whereas application experts designate what is *useful*.

In the given context, this thesis presents image processing methods that transform microscope sensor data (photon counts) to relative dye concentrations of biomolecules of interest across the imaged area.

Sampling and spectra

In image processing, resolution ideally depends on information contained in the measured spatial energy distribution of interest. This section describes sampling criteria for a one-dimensional continuous function $f(x)$, where x is the element of $\mathbb{R}$, i.e., the function f represents variation of the measured physical quantity over space, time, electromagnetic spectrum or similar.

Any regular change of the value of the function $f(x)$ is a sign of pattern. The extent of the change, a frequency ν, should figure as the input variable of a transformation of the input signal $f(x)$ to a new domain. Hence the expression $f(x)\,(\cos(2\pi\nu x) - \imath\sin(2\pi\nu x))$, which gives the value of how much of frequency ν exists in $f(x)$, is the basis of the transformation to the frequency domain:

$$f(x)\,(\cos(2\pi\nu x) - \imath\sin(2\pi\nu x)) = f(x)e^{-2\pi\imath\nu x}. \tag{2.1}$$

The Fourier transform $\mathscr{F}$ integrates the term over the entire domain $\mathbb{R}$

$$\mathscr{F}\{f(x)\} = F(\nu) = \int_{-\infty}^{\infty} f(x)e^{-2\pi\imath\nu x}\,dx. \tag{2.2}$$

The transform exists if the integral of the absolute value of $f(x)$ is finite which is satisfied for all finite functions, i.e., digital signals and images. The Fourier transform is complex, and as such preserves both the magnitude and phase of each frequency ν.

The classical approach is to sample the function f at uniform sampling intervals Δx of the variable x. Mathematically, such sampling function is an impulse train, i.e., sum of periodic impulses Δx units apart, and gives the sampled function

$$\tilde{f}(x) = f(x) \sum_{p_x=-\infty}^{\infty} \delta(x - p_x\Delta x), \tag{2.3}$$

where $p_x \in \mathbb{Z}$. Next, a Fourier transform of the sampled function $\tilde{F}(\nu)$ is derived by means of the Fourier analysis

$$\tilde{F}(\nu) = \frac{1}{\Delta x} \sum_{p_x=-\infty}^{\infty} F\left(\nu - \frac{p_x}{\Delta x}\right). \tag{2.4}$$

Since the aim of sampling is a discrete representation of the input signal or image without loss of information, it is necessary to introduce limitations to $f(x)$. Equation 2.4 suggests that the Fourier transform, as well as the Fourier spectrum, of the sampled function is an infinite periodic sequence of the Fourier transform of the input function $f(x)$ separated $\frac{1}{\Delta x}$ units apart. The input function $f(x)$ can be fully recovered from the sampled function if none of the two contiguous replicates of $F(\nu)$ in $\tilde{F}(\nu)$ overlap. The maximal frequency $\nu_{\max}$ that exists in $f(x)$ should satisfy the condition

$$2\nu_{\max} < \frac{1}{\Delta x}, \tag{2.5}$$

known as the Shannon-Nyquist sampling theorem. The theorem is naturally applicable for digital imaging systems that sample spatial energy distribution in two or even three spatial dimensions.

In practice, reaching the theoretical resolution means that the sampling interval has to be smaller than one-half the period of the finest detail within the image [1]. In biomedical microscopy, that is a never-ending challenge as desire for acquisition of fine details has no limits. Unfortunately for biologists, imaging devices do have resolution limits determined by optics and the wavelength of light [3].

The word spectrum in this section has a common signal processing meaning - the *spectrum* shows how the Fourier transform decomposes signal to its constituent frequencies. Such choice of words was not a coincidence since a few centuries earlier, Newton introduced the same word in science when describing colors dispersed through an optical prism.

2.2 Optical spectroscopy

The purpose of all spectroscopic techniques, from optical, to nuclear magnetic resonance, X-ray, and mass spectroscopy is to give researchers insight into the amount, type or molecular properties of measured materials. This chapter describes principles of optical spectroscopy as well the most available of all spectroscopic techniques - human color vision.

Light and color

In quantum mechanics, according to the wave-particle duality, both light and matter can behave as wave or particle. Light keeps both of its properties while interacting with materials, yet the wave properties of light are primarily of interest in spectral imaging. Measured light is therefore a form of electromagnetic radiation with approximate wavelength range from 400nm to 740m. In addition, the range of the electromagnetic spectrum designated

as light expands significantly when the ultraviolet and infrared parts of the spectrum are included - from nanometres to almost 1mm.

Newton described some spectral properties of light in his famous book published in 1704, *Opticks: Or, A Treatise of the Reflections, Refractions, Inflexions and Colours of Light*. The most interesting experiment, from spectroscopic point of view, is the prism experiment. Newton used a prism to disperse a ray of light to rainbow colors visible on a screen. On the other hand, the book does not contain the drawing of the prism experiment. The engraving in Fig. 2.1 shows Newton himself observing the light spectrum.

Figure 2.1: Engraving of Isaac Newton's prism experiment from 1666. Credit: Science Photo Library, IBL Bildbyrå.

Fig. 2.2 shows how Newton described colors:

> Let GM be produced to X, that MX may be equal to GM, and conceive GX, λX, ιX, ηX, ϵX, γX, αX, MX, to be in proportion to one another, as the numbers, 1, $\frac{8}{9}$, $\frac{5}{6}$, $\frac{3}{4}$, $\frac{2}{3}$, $\frac{3}{5}$, $\frac{9}{16}$, $\frac{1}{2}$, and so to represent the Chords of the Key, and of a Tone, a third Minor, a fourth, a fifth, a sixth Major, a seventh and an eighth above that Key: And the Intervals $M\alpha$, $\alpha\gamma$, $\gamma\epsilon$, $\epsilon\eta$, $\eta\iota$, $\iota\lambda$, and λF, will be the Spaces which the several Colours (red, orange, yellow, green, blue, indigo, violet) take up. [4]

Newton divided the visible part of the spectrum into seven intervals, i.e., primary colors, and used musical tones to set the range of each primary color. Thus red color is between tones with the lowest frequency, it takes two ninths of the visible part of the light spectrum; orange color takes one

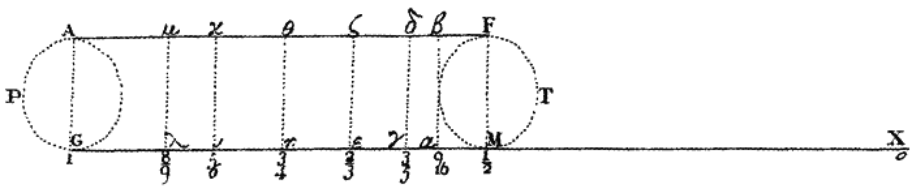


Figure 2.2: Spectrum divided by musical tones. The figure shows the original drawing 4 from Part II of *The First Book of Opticks* [4].

ninth of the visible part of the light spectrum; etc. Even though sound is a mechanical wave and light (as a wave) is electromagnetic radiation, the comparison with musical tones is intuitively correct. Musical tones ideally have line spectra, or at least very narrow spectra, while materials in the nature often emit, reflect or and transmit light characterized by wide spectra.

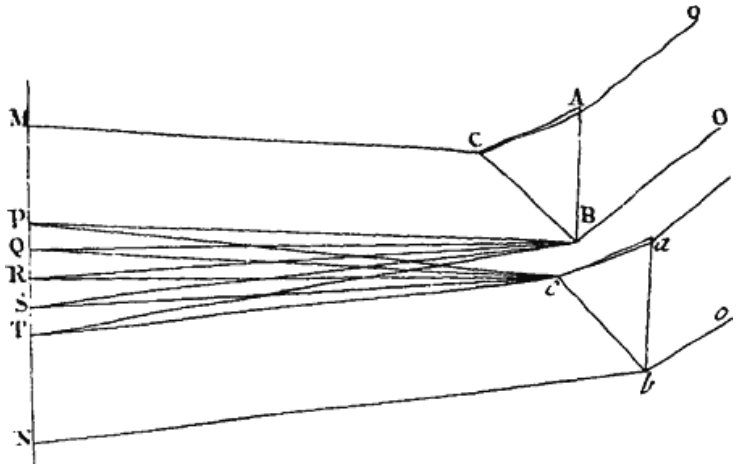


Figure 2.3: Color mixing experiment. Two prisms ABC and abc with equal refracting angles B and b are placed parallel to one another. Light projected through them fall on the screen MN. The figure shows the original drawing 10 from Part II of *The First Book of Opticks* [4].

Newton also practiced mixing colors by using two or more prisms as shown in Fig. 2.3. “And the Colours generated by the interior Limits B and c of the two Prisms, will be mingled at PT, and there compound white”, later followed by conclusion “perfect whiteness may be compounded of Colours” [4]. It is possible to state principles of optical spectroscopy by analyzing the work by Newton. First, each of the colors in the *series of colors* correspond to one of the seven wavelength bandwidths, *spaces*, determined by distinct wavelengths, i.e., *musical tones*. A combination of predefined colors approximate the spectrum of the incident ray, which is a continuous function of wavelength. Next, the spectrum is a line, not a circle, but if a number of *musical tones* with very high and very low

frequencies are *played* simultaneously, the spectrum closes into the circle Fig. 2.4:

> Let the first Part DE represent a red Colour, the second EF orange, the third FG yellow, the fourth CA green, the fifth AB blue, the sixth BC indigo, and the seventh CD violet. And conceive that these are all the Colours of uncompounded Light gradually passing into one another, as they do when made by Prisms; the circumference DEFGABCD, representing the whole Series of Colours from one end of the Sun's colour'd Image to the other, so that from D to E be all degrees of red, at E the mean Colour between red and orange, from E to F all degrees of orange, at F the mean between orange and yellow, from F to G all degrees of yellow, and so on. Let *p* be the center of gravity of the Arch DE, and *q, r, s, t, u, x*, the centers of gravity of the Arches EF, FG, GA, AB, BC and CD respectively, and about those centers of gravity let Circles proportional to the number of Rays of each Colour in the given Mixture be describ'd; that is, the Circle *p* proportional to the number of the red-making Rays in the Mixture, the Circle *q* proportional to the number of the orange-making Rays in the Mixture, and so of the rest. [4]

For instance, a mixture of blue and violet gives a nonspectral color purple, nowadays known as magenta.

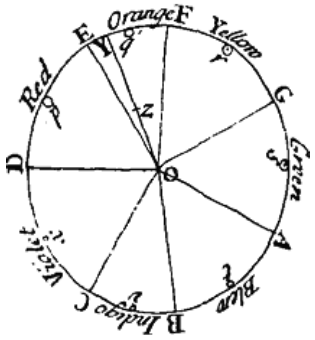


Figure 2.4: Newton's color wheel. The figure shows the original drawing 11 from Part II of *The First Book of Opticks* [4].

There is no doubt that Newton, one of two founding fathers of modern calculus, presented a method for quantitative analysis of light spectra. However, it is important to stress that Newton's papers were not very clear even though he spent more than three decades preparing the book *Opticks*. It also appears as if he did not fully understand that his vision, as well as vision of his colleagues, was trichromatic, and not heptachromatic! For example, in the experiment with two prisms Newton drew conclusions based on appearance of mixed colors:

> For when I was trying this, a Friend coming to visit me, I stopp'd him at the Door, and before I told him what the Colours were, or what I was doing; I

> asked him, Which of the two Whites were the best, and wherein they differed? And after he had at that distance viewed them well, he answer'd, That they were both good Whites, and that he could not say which was best, nor wherein their Colours differed. [4]

Combined with the lack of accurate explanations and drawings, this caused misinterpretations of his work throughout the XIX century [5].

Human color vision vs. a spectroscopic approach

This section explains fundamentals of the human color vision, a well-established field of research in biology, physiology and neuroscience [6]. Eyes detect incident light and generate electro-chemical signals, a role equivalent to that of sensors in electrical engineering. The optic nerve transmits the acquired signal to the image processing unit in the lateral cortex. Finally, the information stream from image based visual sensations combines with visual recollections in the cortex, thus performing image analysis tasks.

In 1980 Bowmaker and Dartnall [7] measured absorption spectra of photoreceptors of human eyes (Fig. 2.5 shows the original figure from the *Journal of Physiology*), photosensitive cells that allow humans to sample spectral properties of observed objects. The photosensitive cones produce signals proportional to the logarithm of the light intensity. Next, retinal ganglio cells combine the signals and generate the electro-chemical output signal consisting of three components:

- luminance – adding the responses of the green and red-sensitive cones
- the red/green ratio – subtracting the responses of the green and red-sensitive cones
- the blue/yellow ratio – subtracting the response of the blue-sensitive cones and the luminance

This way of sampling with three primary colors is rather different in comparison to Newton's quantitative approach where the incident light is a mixture of seven primary colors. Johann Wolfgang von Goethe systematically repeated Newton's experiments and created a new theory in which human color perception is the key concept [8]. Naturally, painters quickly accepted Goethe's standpoint and also influenced the modern color theory. For example, modern image compression algorithms reduce the amount of data while changes in the appearance of an image should remain below just-noticeable difference. Another byproduct of Goethe's color theory is inclusion of extraspectral colors in the color wheel as red-blue mixtures of primary colors appeared as relevant as cyan and yellow, all three denoted as secondary colors [9].

Image processing literature often prioritizes perceived color instead of offering Newton's quantitative approach. A typical example is transforma-

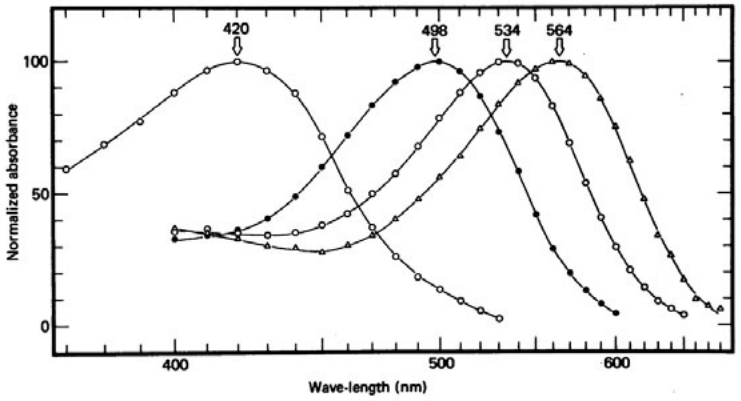


Figure 2.5: Primates have four types of photoreceptors in the retina of an eye, each absorbing light over a wide range of wavelengths [7]. The three curves labeled with 420, 534 and 564 are mean absorbance spectra of blue, green and red-sensitive cones, respectively. They are active in well illuminated surroundings. The curve with the peak at 498 nm is the mean absorbance spectrum of the rods, photoreceptors active in dark surrounding. Reprinted with the permission from John Wiley & Sons Ltd.

tion of input red-green-blue triplets to hue-saturation-luminance [9]. *"Hue represents dominant color as perceived by an observer. Thus when we call an object red, orange, or yellow, we are referring to its hue. Saturation refers to the relative purity or the amount of white light mixed with a hue."* Therefore, in color image processing, even the non-spectral magenta which is a balanced red-blue mixture, is considered as a pure spectral color. Since magenta is not associated with any wavelengths, it is then unclear why a perfectly balanced red-green-blue mixture (the gray color) is left out. To conclude, with respect to quantitative analysis of spectral information, Newton's approach is unambiguous.

Interestingly, not all animals have trichromatic vision. Mantis shrimps use 16 types of photoreceptors with twelve different absorbance spectra [10]. So far, no one has described an animal species with seven types of photoreceptors, each corresponding to distinct wavelength bandwidths – just like Newton's primary colors. Irrespectively of the number of types of photoreceptors n_c, eyes are sensors that produce ordered n_c-tuples as they unevenly sample the light spectrum and provide insight into the nature of the observed object.

Material investigation

Optical spectroscopic techniques facilitate analysis of light-material interactions by using an external light source that illuminates the material [11].

In emission spectroscopy, the material partially absorbs and re-emits the incident light in all directions. Emission spectroscopy analyzes the spectrum of the re-emitted light; usually different from the spectrum of the incident light. Another type of spectroscopy is absorption spectroscopy. In this experimental setup the incident light beam is attenuated by the material, hence the difference between the incident and transmitted light exhibits material properties.

Modern spectrophotometers resemble Newton's prism experiment, sometimes with one addition: instead of optical prisms, diffraction gratings are preferred (Fig. 2.6). This approach allows decomposition of the incident beam into a number of spatially separated spectral components, i.e., samples of the continuous spectral profile of the beam. Therefore, from an engineering point of view, the challenge is to design a sensor that counts the number of incident photons to electric signals and provides a sampled spectrum.

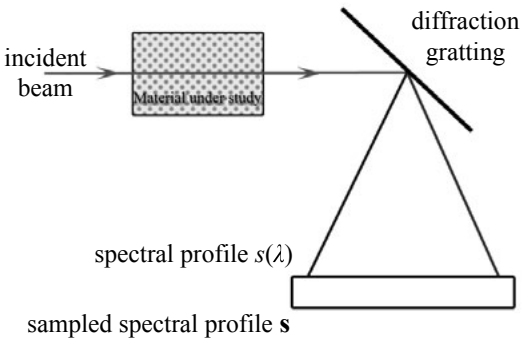


Figure 2.6: Absorption spectroscopy may be described in terms of the Newton's original theory. The incident *achromatic light* beam passes through the material and the grating diffracts the transmitted light which then falls onto the sensor.

2.3 Optical imaging systems

Optical imaging systems provide information about radiant energy reflected or emitted from the material, or transmitted through the material at all spatial position $\mathbf{p}$ and wavelength λ [1]. The level of detail provided by an ideal sensor, i.e., resolution of such multidimensional image $s_{\mathbf{p}}(\lambda)$, depends on conditions determined by the Shannon-Nyquist theorem. The same theorem holds even if the sensor acquires images at all times.

However, real optical imaging systems have a number of limitations. In general, every system imposes restriction on the maximum intensity, i.e.,

saturation level of the sensor s^{sat}

$$0 \leqslant s_{\mathbf{p}}(\lambda) \leqslant s^{sat}, \tag{2.6}$$

as well as limitations of the range of spatial positions where the image is captured. The first step is sampling the light spectrum using a number of optical elements. The imaging system gives the sampled spectrum, i.e., the spectral image:

$$s_{\mathbf{p},k} = \int_{\lambda} R(\lambda) F_k(\lambda) s_{\mathbf{p}}(\lambda) d\lambda, \tag{2.7}$$

where $F_k(\lambda), k = 1, ..., n_c$, are spectral responses of optical filters and $R(\lambda)$ is the transfer function of the sensor. In practice, sampling of spectral information is implemented by compromising the level of spatial and temporal details [12]. For example, the system shown in Fig. 2.6 provides spectral information at the cost of one spatial dimension. If the spectral resolution is 10nm or less, the system collects on the order of ten or more spectral channels. If the maximum frequency in the signal along the λ is band limited by the grating, this system satisfies the Shannon-Nyquist theorem. But as the following chapters describe, it is very common to undersample the light spectrum using optical filters that transmit light over wide wavelength bandwidths. Figure 2.7 shows spectral responses $F_k(\lambda)$.

The most common imaging sensors today are two-dimensional sensor arrays that uniformly sample spatial information, e.g., charge-coupled devices (CCD) and their low-cost alternatives based on complementary metaloxide-semiconductors (CMOS). Unlike the human eye photoreceptors with logarithmic transfer function, the CCD sensors have linear transfer function (Fig. 2.8), i.e., the output signal is proportional to the number of photons in the incident light [13, 14].

Limitations of optical elements primarily affect sampling of spatial and spectral information. In addition, optical imaging sensors have a number of limitations with respect to dynamic range of the output signals. During photon production, the number of photons emitted from a constant light source over a finite time interval is stochastic [15]. Under normal operating conditions, it is Poisson distributed and represents the dominant source of noise in sensors, denoted as the **photon noise**.

The dynamic range of sensors is defined as a ratio between the maximal and minimal measurable values. In the process of converting the number of incident photons to the digital output, a number of physical processes affect the noise level and consequently dynamic range of the system and the signal-to-noise ratio [13, 14, 15, 16, 17]:

- Incident light generates electron-hole pairs separated by an electric field. It exists due to uncertainty of the number of generated electrons and, just like photon noise, follows Poisson distribution.

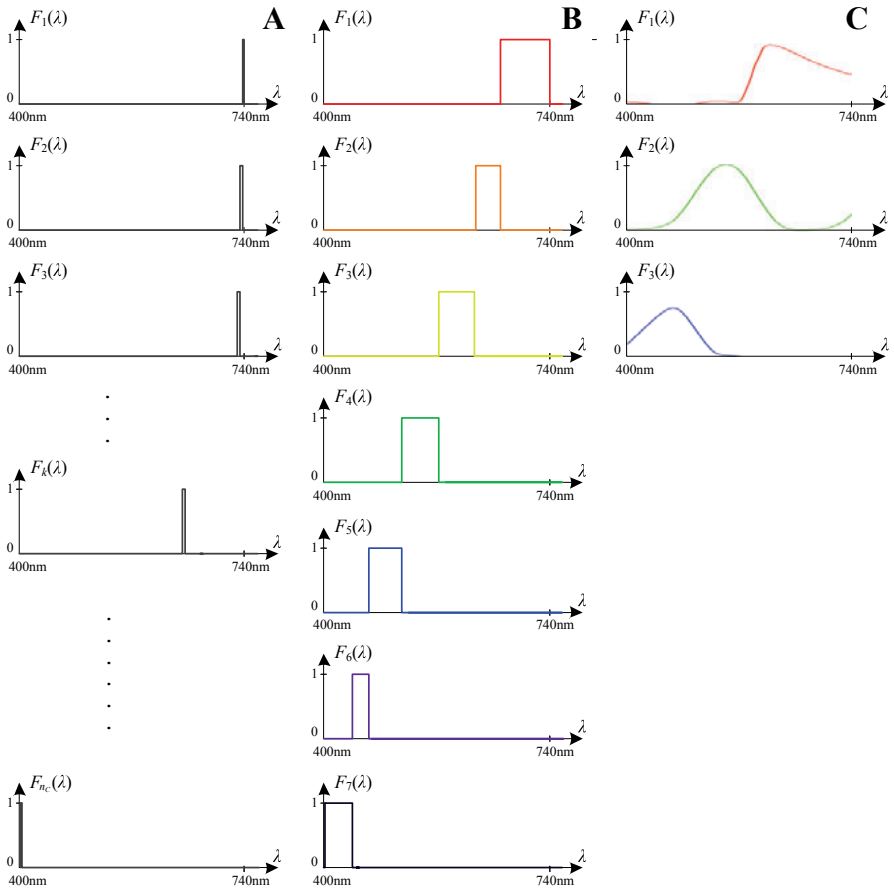


Figure 2.7: Examples of spectral responses of three common optical imaging systems. (A) A hyperspectral camera uniformly samples the spectrum, $n_C \sim 10$ or even $n_C > 100$ spectral channels. (B) A multispectral camera irregularly samples the spectrum. This example illustrates the cut-off frequencies determined by Newton's musical tones shown in Fig. 2.2. (C) A standard tri-color RGB camera. Optical filters are modeled by human photoreceptor absorbance spectra. Note that (A) and (B) show idealized bandpass filters. In practice, there are transition zones from full transmission to full blocking depending on how the filters are implemented.

- Some electrons are thermally generated irrespectively of the number of incident photons. Their number increases with temperature and results in **dark current noise**.
- The electron flow from the semiconductor accumulates on the positive plate of the capacitors at the input of the operational amplifier circuits. Amplifiers integrate the number of electrons and produce output voltage. **Readout noise** models fluctuations from the linear transfer function of amplifier circuits, and it is dominant only when amplifiers are read

at high rate. In addition, **impulse noise** (also known as salt-and-pepper noise in image processing literature [9]) used to be a relevant source of noise up to the late 1970's [16]. It was caused by malfunction of early operational amplifiers.

- **Quantization noise** is present even in ideal noise-free sensors. It is a result of converting the measured analogue value to the discrete domain. The number of quantization levels should be chosen to allow maximal dynamic range of the sensor.

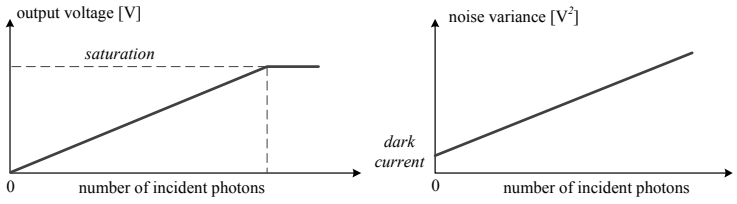


Figure 2.8: Important properties of CCD sensors – both the output measurements and noise variance are linearly dependent on the number of incident photons. The consequence of which is that signal-to-noise ratio grows with the square root of the signal.

A previous section describes signal and image processing as a group of techniques that aim to represent and transform measurements from sensors using some *useful* operation. In optical imaging, examples of *useful* operations are suppression of noise by filtering methods or compensation for distortions introduced by the imaging system by deconvolution [3]. But this is often just an initial step in high-level image processing techniques.

2.4 Imaging in natural sciences and medicine

Information processing from image based measurements

Today, imaging techniques play a prominent role in materials science, earth sciences, chemistry and biology as well as in many fields of medicine, particularly in radiology and pathology. This section describes the significance of digital image processing techniques in context of the generalized information processing flowchart shown in Fig. 2.9.

Scientific problem. Medical researchers employ scientific methods to answer questions about a cause of a disease and treatment or provide medical doctors with diagnostic tools. For instance, in radiology, one strives to provide information about internal organs and tissues without harming the patient – a mission impossible to accomplish without use of imaging. Not far from radiology (at least from engineering point of view), pathologists

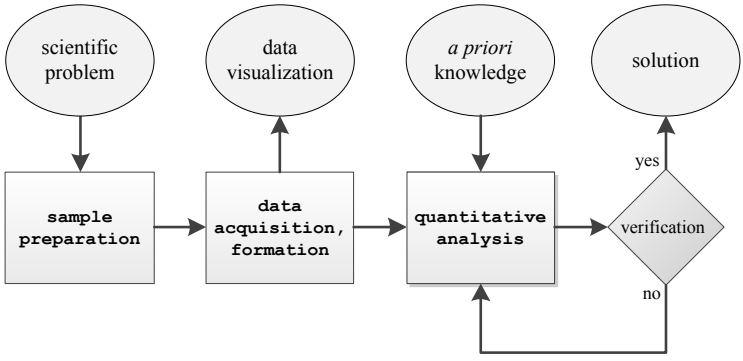


Figure 2.9: Information processing in sciences and medicine, from problem to solution. Image processing is closely tied to data acquisition and quantitative analysis, helping experts to find a solution.

provide information about the mechanism of a disease, from cause to manifestation and, in particular, look for changes in tissues and organs, i.e., a manifestation of a disease or injury. Describing content of a tissue biopsy in a quantitative manner is a typical problem addressed in pathology and employed for malignancy grading. On the other hand, in sciences, researchers pose hypotheses and need to either confirm or disprove them. Both outcomes are equally favorable **solutions** to the problem in the context of testing the hypothesis.

Specimen preparation as well as patient preparation for radiological procedures comprises a number of methods employed to enhance the measurable physical quantity of interest. For instance, in X-ray based imaging techniques, diatrizoic acid acts as a contrast agent which amplifies the signal from blood vessels, while in fluorescence microscopy, 4',6-diamidino-2-phenylindole binds and labels primarily DNA. This step is particularly important for spectral image processing, and is therefore addressed in more details in subsequent section.

Data acquisition is sampling of spatial energy distribution radiated from the specimen and storing discretized values on a digital storage medium. Acquisition is, as the previous section describes, closely related to image formation, a process that transforms the data to a form suitable for processing as well as visualization on an arbitrary display.

Quantitative analysis may be based on mathematical modeling, machine learning, or a combination of the two. It often starts with image segmentation, a method for separating individual objects from the background, where an object is a spatially connected set of imaged

positions. In approaches based on machine learning, a priori knowledge consists of a number of numerical descriptors assigned to individual objects, e.g., object size or shape. Numerical descriptors, so called features, build feature vectors that span multidimensional spaces. A training set of images, where the solution is known, determines what features are statistically significant, and provides classification rules. During the verification procedure, consistency of the classification rules is tested.

While a priori knowledge of underlying processes in the machine learning approach provides a set of possibly relevant features, it is essential for mathematical modeling. Basic laws of physics, chemistry or biology describe the energy radiated from the specimen in form of a system of equations. Statistical methods may be employed to estimate parameters of the model. Once the analytical model and its limitations are established, this approach does not require verification as the solution of the system of equations is the solution to the problem.

The following section describes staining methods used in biotechnology and pathology for identification of biomolecules of interest.

Staining – from histological labeling to detection of single molecule

In optical imaging, tissues and cells exhibit limited spectral characteristics and display limited contrast between different imaged positions. Detection of different cellular, subcellular and molecular structures or events *in situ* is possible with spatial resolution that satisfies the condition shown in eq. 2.5 and enhancement by an appropriate staining method. Staining helps visualizing and quantifying desired subcellular entities required for identification, while no dramatic changes disrupt the native cell or tissue morphology [18].

Histochemistry favors selective staining with chemical compounds (dyes) that specifically interact with particular cellular components. For instance, a commonly used histochemical dye hematoxylin specifically stains the cell nucleus while eosin stains the cytoplasm and connective tissue. A wide range of dyes is available for the staining of different cellular structures providing distinct spectral signatures. However, not every cellular component is large enough or sufficiently abundant in every tissue to be detected by a histochemistry approach. In such situations, the small amount of stained biomolecules (signal) does not introduce detectable changes in characteristic spectra.

Here immunohistochemistry techniques can be applied to amplify biochemical signals by increasing local concentrations of dyes, and consequently yielding higher biochemical signal-to-background ratio. The method employs antibodies that specifically label proteins. In optical imaging, the location of an antibody is detected by binding a dye molecule

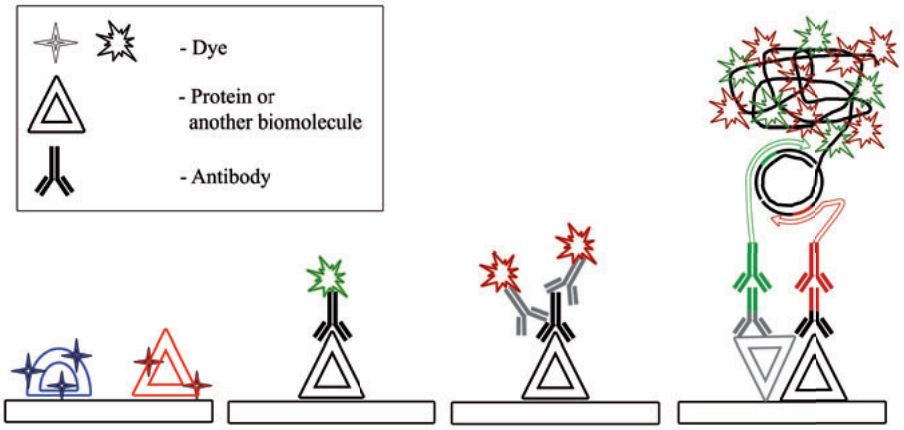


Figure 2.10: Overview of labeling techniques, from left to right: histochemistry, direct immunostaining, indirect immunostaining, detection of molecule interactions by rolling circle amplification [19].

to the antibody. This is the direct method of immunostaining. If the signal obtained from direct immunostaining remains insufficiently strong to be detected, one relies instead on indirect immunostaining, a method that implies usage of a dye-labeled secondary antibody that binds specifically to the primary antibody bound to the target. This which ultimately results in increased biochemical signal [18].

There are still ways to amplify the signal derived from the secondary antibodies, with additional modifications. Having such properties, immunohistochemistry is an invaluable technique not only for establishing the presence of cellular and molecular targets but also for determining their spatial and temporal localization. The detection of biomolecular structure and function extends itself even further by making targets accessible to the antibody while maintaining morphology of the specimen intact. In reality there is a number of methods for single molecule detection and a handful of methods and techniques available for the detection of protein-protein interactions and specific DNA sequences *in situ*. But the majority of presently available methods for detecting protein-protein interactions assume altering the proteins in question in such way that their function can be potentially disrupted [20]. In contrast, there is a growing demand for methods for detection of single molecules and single molecule interactions in their native environments. Here one cannot rely on histochemistry since high concentrations of dye molecules interact unspecifically with the tissue [21, 22], causing high levels of biochemical noise.

Methods which fulfill such demands for single molecule detection are padlock probes [23] which target a specific DNA segment. Padlock probes hybridize with the DNA segment, get ligated, and replicated by rolling cir-

cle amplification. Similarly, rolling circle amplification can amplify the signal originating from proteins or protein-protein interactions detected by proximity ligation [19]. In addition to specific detection, rolling circle amplification as well as histochemical and immunohistochemical techniques fulfill the requirements of the model described in the following chapter.

3. Linear mixture model

Essentially, all models are wrong, but some are useful.
– George Edward Pelham Box (born 1919)

The linear mixture model has been widely accepted in the microscopy community [24]. The model assumes that the spectral signature $c_{\mathbf{p}}(\lambda)$ of dyes mixed at imaged positions $\mathbf{p}$ is a linear combination of spectral signatures of individual dyes. In addition, the model makes the assumption that the contribution of each individual dye is proportional to its molar concentration. The aim of methods based on the linear mixture model is to estimate relative concentrations of individual dyes $\mathbf{d_p}$ from the spectral image data. This is known as color compensation [25, 26], color deconvolution [27], unmixing [28, 29] or decomposition (Papers I and IV).

3.1 Fluorescence microscopy

In both wide-field and confocal fluorescence microscopy, optical elements such as filters or grating sample light spectra and acquire spectral images [3, 12]. Fig. 3.1 shows how the microscope uses a light source to excite fluorescent dyes and then acquires light emitted from the specimen. The contribution of dye j to spectral channel k at imaged position $\mathbf{p}$ is [30, 31]

$$c_{\mathbf{p},k,j} = t_k l_{\mathbf{p}} \varsigma_{\mathbf{p}} \rho_{\mathbf{p},j} \int_{\lambda} R(\lambda) T_{em,\mathbf{p},j}(\lambda) F_{em,k}(\lambda) d\lambda \int_{\lambda'} L(\lambda') T_{exc,\mathbf{p},j}(\lambda') F_{exc,k}(\lambda') d\lambda', \tag{3.1}$$

where

- for each imaged position $\rho_{\mathbf{p},j}$ is the molar concentration of dye j, $\mathbf{p}$, $l_{\mathbf{p}}$ is optical depth (the measure of the fraction of photons emitted from the specimen that fall onto the sensor) and $\varsigma_{\mathbf{p}}$ is the area covered by the sensor element,
- $R(\lambda)$ is the transfer function of the CCD sensor, e.g., quantum efficiency, a probability that a photon of wavelength λ hitting the sensor generates an electron-hole pair,
- $L(\lambda)$ is the excitation source flux at wavelength λ,
- for each fluorescent dye j at imaged position $\mathbf{p}$, $T_{exc,\mathbf{p},j}(\lambda)$ and $T_{em,\mathbf{p},j}(\lambda)$ are excitation and emission spectra, respectively, i.e., probabilities of incident photon absorption and emission, at certain wavelength λ,

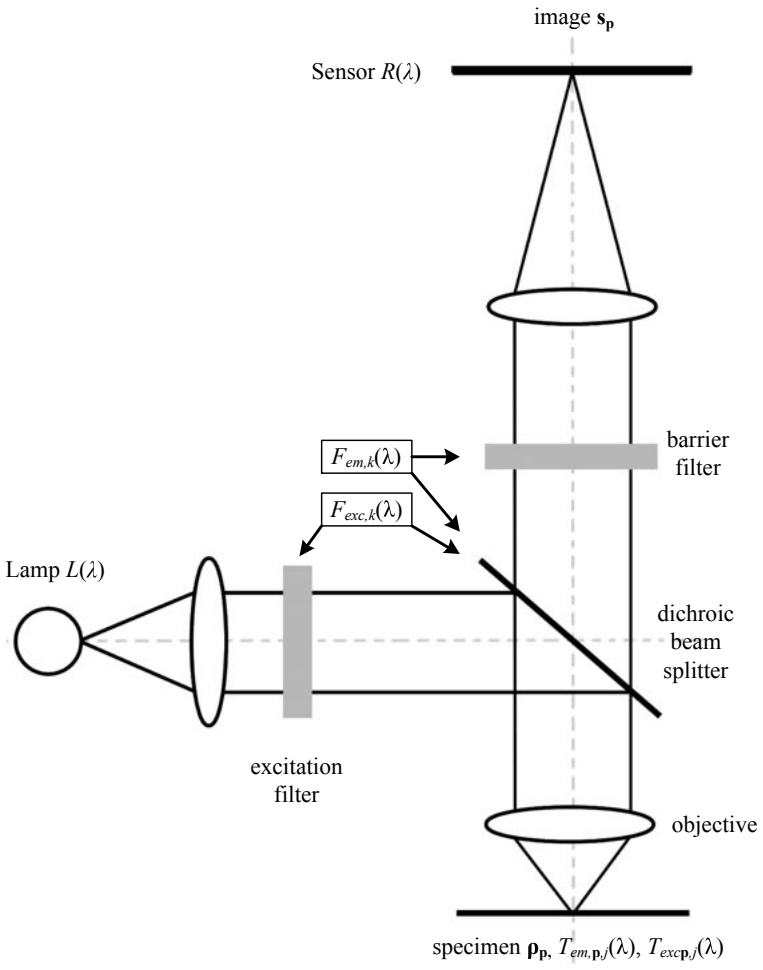


Figure 3.1: Simplified diagram of a fluorescence microscope.

- for imaging in the channel k, t_k is the exposure time, $F_{exc,k}(\lambda)$ are combined transmission spectra of the excitation filters and reflection spectra of dichroic mirrors, $F_{em,k}(\lambda)$ are combined transmission spectra of dichroic mirrors and barrier filters.

Usually sensor elements have the same area and the optical depth does not vary much over the specimen, thus $\forall \mathbf{p}, \varsigma_\mathbf{p} \equiv \varsigma, l_\mathbf{p} \equiv l$. In spite of the fact that spectral imaging systems acquire only emitted photons the product suggests that emission is dependent on excitation. According to the linear model, the total measured light intensity is

$$s_{\mathbf{p},k} = \sum_{j=1}^{n} c_{\mathbf{p},k,j} + b_{\mathbf{p},k}, \tag{3.2}$$

where $b_{\mathbf{p},k}$ is the black level offset for channel k at imaged position $\mathbf{p}$ (also dependent on the exposure time). By substituting equation 3.1, equation 3.2 becomes

$$s_{\mathbf{p},k} = \sum_{j=1}^{n} \underbrace{t_k l \varsigma \int_{\lambda} R(\lambda) T_{em,\mathbf{p},j}(\lambda) F_{em,k}(\lambda) d\lambda \int_{\lambda'} L(\lambda') T_{exc,\mathbf{p},j}(\lambda') F_{exc,k}(\lambda') d\lambda'}_{a_{\mathbf{p},k,j}} \rho_{\mathbf{p},j} + b_{\mathbf{p},k}. \tag{3.3}$$

The equation shows that the transfer function of the system $a_{\mathbf{p},k,j}$ is not constant due to variations in probabilities of photon absorption and emission over the specimen, i.e., excitation and emission spectra,

$$s_{\mathbf{p},k} = \sum_{j=1}^{n} a_{\mathbf{p},k,j} \rho_{\mathbf{p},j} + b_{\mathbf{p},k}. \tag{3.4}$$

Assuming that for each j, the differences in spectra over the specimen are never greater than variations in spectra between the dye j and any other dye, the equation 3.4 can be written as

$$s_{\mathbf{p},k} = \sum_{j=1}^{n} a_{k,j} \rho_{\mathbf{p},j} + b_{\mathbf{p},k} + \epsilon_{\mathbf{p},k}, \tag{3.5}$$

where $a_{k,j}$ depends on the average values of $T_{exc,\mathbf{p},j}(\lambda)$ and $T_{em,\mathbf{p},j}(\lambda)$ over all imaged positions $\mathbf{p}$ and $\epsilon_{\mathbf{p},k}$ is the biochemical noise term, the measure of variation in dye spectra over the specimen [31]. The background level is assumed to be equal to the black level offset, hence the noise term is zero-mean and random after background subtraction.

If the parameters of the transfer function $a_{k,j}$ are known, the system of linear equations 3.5 allows estimation of molar concentration $\rho_{\mathbf{p},j}$ up to a constant, i.e., the relative dye concentration $d_{\mathbf{p},j} \propto \hat{\rho}_{\mathbf{p},j}$ yielded by minimizing the noise term

$$s_{\mathbf{p},k} - b_{\mathbf{p},k} = \sum_{j=1}^{n} a_{k,j} d_{\mathbf{p},j}, \tag{3.6}$$

which can be written in vector form as

$$\mathbf{s_p} - \mathbf{b_p} = \mathbf{A}\mathbf{d_p}, \tag{3.7}$$

where the columns of the mixing matrix $\mathbf{A}$ are sampled spectral signatures of respective dyes in the mixture.

The sampled spectral signature of the mixture $\mathbf{c_p}$ is thus in linear relationship with the spectral image $\mathbf{s_p}$ and, as stated above, $\mathbf{d_p}$ estimates $\rho_{\mathbf{p}}$ up to a constant. For the sake of simplicity, the 1-norm of sampled spectral

signature of a dye $\mathbf{a}_j$ is arbitrarily set to one. Therefore,

$$\mathbf{c_p} = \mathbf{A}\mathbf{d_p}, \tag{3.8}$$

where $\mathbf{A}$ is the mixing matrix – it determines this model which is linear in its parameters.

3.2 Bright-field microscopy

From a spectroscopic point of view, the absorption spectroscopy experimental setup in Fig. 2.6 may be considered as a single-pixel bright-field microscope. Unlike fluorescence microscopy, in bright-field microscopy the relationship between spectral images and sampled spectral signatures of mixtures is non-linear and requires application of the Beer-Lambert law of absorption in order to linearize the mixture [24, 27, 32].

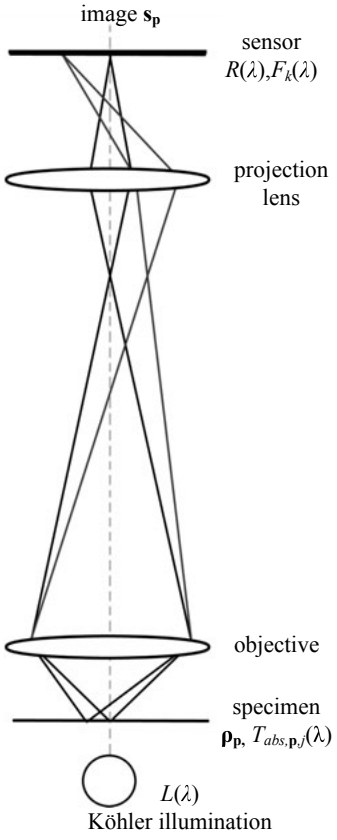


Figure 3.2: Simplified diagram of a bright-field microscope.

Fig. 3.2 shows the imaging system with the following non-linear transfer function:

$$s_{\mathbf{p},k} = b_k + t w_k l_{\mathbf{p}} \varsigma_{\mathbf{p}} \int_\lambda L(\lambda) \exp\left(- \sum_{j=1}^{n} T_{abs,\mathbf{p},j}(\lambda) \rho_{\mathbf{p},j} \right) F_k(\lambda) R(\lambda) d\lambda, \quad (3.9)$$

where

- for each imaged position, $\rho_{\mathbf{p},j}$ is molar concentration of the dye j, $\mathbf{p}$, $l_{\mathbf{p}}$ is optical depth and $\varsigma_{\mathbf{p}}$ is the area covered by the sensor element, $\forall \mathbf{p}, \varsigma_{\mathbf{p}} \equiv \varsigma, l_{\mathbf{p}} \equiv l$,
- t is the exposure time and $R(\lambda)$ is the transfer function of the CCD sensor,
- b_k is the background offset and $L(\lambda)$ is the the light source flux at wavelength λ; it is not dependent on $\mathbf{p}$ as Köhler illumination corrects for otherwise non-uniform illumination from the lamp,
- for each fluorescent dye j at imaged position $\mathbf{p}$, $T_{abs,\mathbf{p},j}(\lambda)$ is the absorbance at certain wavelength λ,
- for imaging in the channel k, w_k is the operational amplifier gain, $F_k(\lambda)$ are spectral responses of optical filters.

In fluorescence microscopy systems (Section 3.1), spectral channels are acquired sequentially, while bright-field microscopes acquire all spectral channels simultaneously, thus $\forall k_1, k_2 = 1, ..., n_c, t_{k_1} = t_{k_1} \equiv t$. In addition, the background offset can be easily set to zero, $\forall k = 1, ..., n_c, b_k = 0$. The system response to the blank image $\mathbf{s}^0$ is recorded

$$\forall \mathbf{p}, j, \rho_{\mathbf{p},j} \Rightarrow s_k^0 = t w_k l \varsigma \int_\lambda L(\lambda) F_k(\lambda) R(\lambda) d\lambda. \quad (3.10)$$

The operational amplifier gains w_k tune the spectral signature of the blank image $\mathbf{s}^0$. For instance, if the microscope is equipped with an RGB camera setting the gains to achieve $s_1^0 = s_2^0 = s_3^0$ is referred to as white balance [33].

According to the first mean value theorem for integration applied n times to equation 3.9, there exist n wavelengths for each spectral channel such that

$$s_{\mathbf{p},k} = t w_k l \varsigma \prod_{j=1}^{n} \exp\left(-T_{abs,\mathbf{p},j}(\lambda_{j,k}) \rho_{\mathbf{p},j}\right) \int_\lambda L(\lambda) F_k(\lambda) R(\lambda) d\lambda, \quad (3.11)$$

which by substituting equation 3.10, becomes:

$$s_{\mathbf{p},k} = s_k^0 \exp\left(- \sum_{j=1}^{n} T_{abs,\mathbf{p},j}(\lambda_{j,k}) \rho_{\mathbf{p},j} \right). \quad (3.12)$$

By applying the logarithm to each side of the equation the model linear in its parameters for bright-field microscope is derived:

$$\ln s_k^0 - \ln s_{\mathbf{p},k} = \sum_{j=1}^{n} T_{abs,\mathbf{p},j}(\lambda_{j,k})\rho_{\mathbf{p},j}. \tag{3.13}$$

Analogously to the fluorescence microscope system, here the biochemical noise term originates from varying absorbance over the specimen. Following the same procedure, for each j, relative dye concentration $d_{\mathbf{p},j}$ estimates molar concentrations $\rho_{\mathbf{p},j}$ up to a scalar constant, so that the biochemical noise term is minimal, thus:

$$\ln s_k^0 - \ln s_{\mathbf{p},k} = \sum_{j=1}^{n} a_{k,j} d_{\mathbf{p},j}. \tag{3.14}$$

By introducing sampled spectral signature of a mixture $\mathbf{c_p}$ to the model, $\forall k = 1, ..., n_c, c_{\mathbf{p},k} \equiv \ln s_k^0 - \ln s_{\mathbf{p},k}$, the previous equation is transformed to:

$$c_{\mathbf{p},k} = \sum_{j=1}^{n} a_{k,j} d_{\mathbf{p},j}, \tag{3.15}$$

which can be written in vector form:

$$\mathbf{c_p} = \mathbf{A}\mathbf{d_p}. \tag{3.16}$$

3.3 Parameter estimation

Least-squares fit

Equations 3.8 and 3.16 show that applying matrix pseudo-inversion results in relative dye concentrations. Naturally, sampled spectral signatures of dyes $\mathbf{a}_j$ that represent columns of the mixing matrix $\mathbf{A}$ need to be determined beforehand. For instance, Dickinson *et al.* [34] and Ruifrok [27] published the following simple procedure in fluorescence and bright-field microscopy, respectively. They recorded sampled spectral signatures of pure dyes, normalized them to the unit length, and used them as approximation of the mixing matrix.

The procedure requires that the mixing matrix has full rank, i.e., its columns are linearly independent. If $n = n_c$, relative dye concentrations at imaged positions $\mathbf{p}$ are estimated by the unique solution of the least square problem [35]:

$$\mathbf{d_p} = \mathbf{A}^{-1}\mathbf{c_p}, \tag{3.17}$$

or if the number of spectral channels is greater using the matrix pseudo-inversion:

$$\mathbf{c_p} = \mathbf{A}\mathbf{d_p} \Rightarrow \mathbf{A}^T\mathbf{c_p} = \mathbf{A}^T\mathbf{A}\mathbf{d_p} \Rightarrow \mathbf{d_p} = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{c_p}, \tag{3.18}$$

which is the optimal solution of overdetermined least-squares problems [35]. In addition, if the number of spectral channels is greater than n, singular value decomposition guarantees numerically stable solutions.

Estimation of unknown model parameters by optimization

An alternative to the manual approach is using experimental data to estimate elements of the mixing matrix. Matrix $\mathbf{C}$, $\mathbf{C} \in \mathbb{R}^{n_c \times n_\mathbf{p}}$, stores $n_\mathbf{p}$ recorded measurements $\mathbf{c_p}$. Result of the algorithm is the mixing matrix $\mathbf{A}$ and estimates of relative dye concentrations $\mathbf{D}$, $\mathbf{D} \in \mathbb{R}^{n \times n_\mathbf{p}}$ yielded by solving an optimization problem.

$$\text{Minimize } \|\mathbf{C} - \mathbf{A}\mathbf{D}\|^2 \text{ subject to } (\forall j,k)\, a_{k,j} \geqslant 0\ (\forall \mathbf{p}, j)\, d_{\mathbf{p},j} \geqslant 0, \tag{3.19}$$

where $\|\cdot\|^2$ is the sum of squares of the elements of the mixing matrix. Non-negative matrix factorization [36] is a common name for this optimization problem.

Methods described in this section first found application in processing of hyperspectral satellite imagery [37], when $n_c \gg n$ and the spectra are sampled uniformly. The following chapter presents a method for parameter estimation and linear decomposition adapted for biomedical applications. Limitations of the linear mixture model are presented in Chapter 6.

4. Methods for decoupling light intensity and spectral information

Prediction is very difficult, especially about the future.
– Niels Bohr (1885 - 1962)

The previous chapter described the model linear in parameters which allows estimation of relative dye concentrations. The model parameters are linearly independent sampled spectral signatures of dyes. If the parameters are known, the solution to the least squares problem is the estimate of relative dye concentrations. If the model parameters are unknown, non-negative matrix factorization can be employed to the training data to estimate them, but other estimation techniques may also be of interest as well [29, 38]. This chapter presents a method for estimation of the model parameters that covers a wider range of light microscopy applications where two or three spectral channels are used to provide spectral information.

4.1 Noise compensation

Fig. 4.1A shows a histological specimen stained with hematoxylin. It is a tri-channel spectral image acquired using a bright-field microscope and optical filters with spectral responses illustrated in Figure 2.7C. Figure 4.1C shows a scatter plot of the distribution of the image data $[s_{\mathbf{p},1}\, s_{\mathbf{p},2}\, s_{\mathbf{p},3}]^T$ in a three-dimensional color space. The data is distributed from $\mathbf{s}^0$, i.e., white regions with low molar dye concentration, to gray-blue and bends to the dark blue. Clearly, the main source of variation in the specimen is the dye concentration that varies with light intensity. Fig. 4.1C also shows the biochemical noise described in the previous chapter, i.e., data points associated with the same dye hematoxylin, on the same distance from $\mathbf{s}^0$ exhibit different spectral properties. Figure 4.1B shows the specimen imaged after eosin was added. The counter-stain eosin also affects the color of hematoxylin stained tissue as illustrated in Fig. 4.1D.

While a general model for characterization of biochemical noise has not been developed, CCD sensor noise is described in Section 2.3. Sensor noise disturbs regions with low molar concentration and a common approach is to set an arbitrary background intensity threshold. However, in addition to

introducing user bias, this approach also removes regions with low dye concentration. Fig. 4.2 shows a procedure for noise estimation, i.e., a step performed prior to image acquisition. Paper I presents a method for quantization noise compensation in ideal sensors and Paper IV introduces a model for photon noise in CCD sensors.

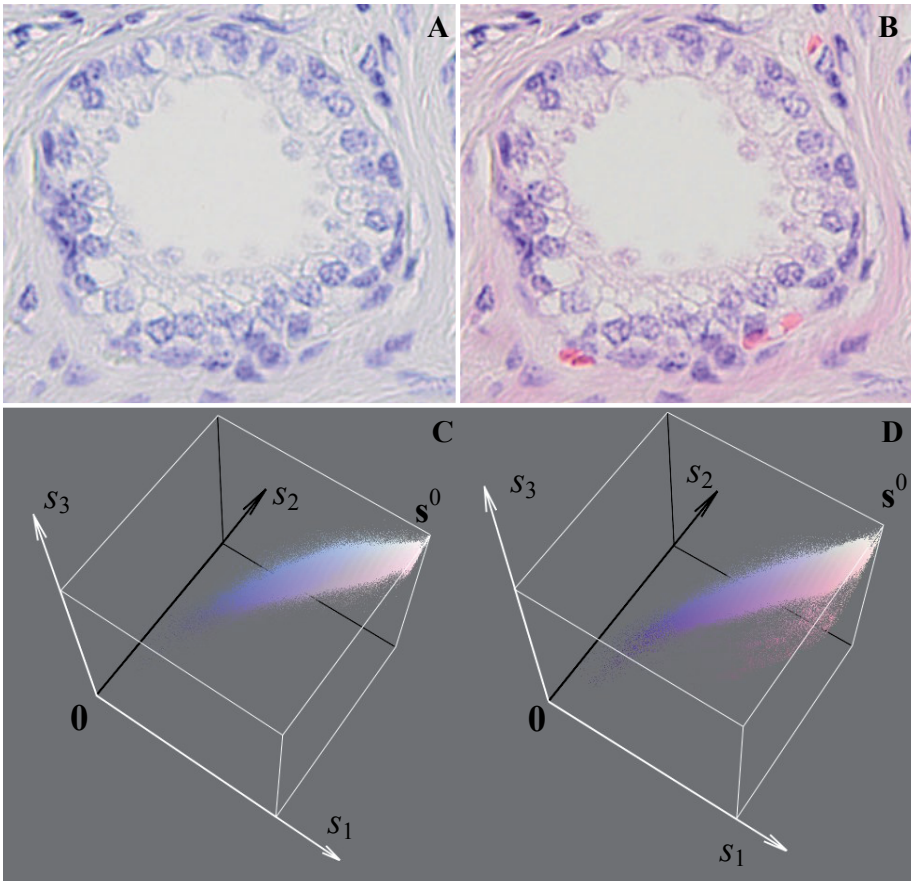


Figure 4.1: Hematoxylin (A) and Hematoxylin/Eosin (B) stained prostate gland section and corresponding scatter plots (C,D).

4.2 Chromaticity spaces

The purpose of chromaticity spaces is to present spectral information and remove intensity variations. In spectral images, intensity is approximated by the length of the sampled spectral signature of the mixture **c**:

$$\|\mathbf{c}\|_1 = \sum_{k=1}^{n} c_k \approx \int_{\lambda} c(\lambda) d\lambda. \tag{4.1}$$

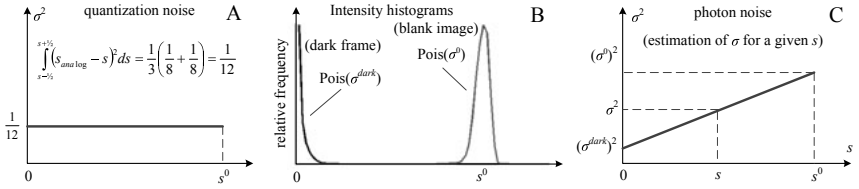


Figure 4.2: An ideal sensor (A) is dominated by quantization noise. In modern CCD sensors, however, the level of dark current noise determines the dynamic range of the sensor and consequently the number of quantization levels. Then, the photon noise is dominant (B) and can be estimated by measuring variance of the dark frame and the blank image. As the photon noise is Poisson distributed, standard deviation grows with the square root of the signal (C).

For each imaged position $\mathbf{p}$, the sampled spectral signature of the mixture $\mathbf{c_p}$ is either directly proportional to the spectral image $\mathbf{s_p}$ compensated for the background fluorescence, or it requires transformation by the Beer-Lambert law of absorption. After the transformation to a chromaticity space, two spectral signatures with different intensities but identical shape of the corresponding continuous spectral signatures overlap in the same position within the plane.

Paper IV implements decoupling as perspective projection to the chromaticity plane of the sampled spectral signature of the mixture [39]:

$$\begin{bmatrix} \alpha_{\mathbf{p}} \\ \beta_{\mathbf{p}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \sqrt{\frac{2}{3}} \end{bmatrix} \frac{\mathbf{c_p}}{\|\mathbf{c_p}\|_1}. \tag{4.2}$$

As an alternative, Paper I presents a method for quantization noise compensation incorporated in decoupling of intensity and spectral information. For each imaged position $\mathbf{p}$, sampled spectral signature $\mathbf{c_p}$ is presented as a ratio of the two spectral channels, i.e., an angle

$$\alpha_{\mathbf{p}} = \arctan \frac{c_{\mathbf{p},2}}{c_{\mathbf{p},1}}. \tag{4.3}$$

Analysis of patterns in chromaticity spaces

Figure 4.3 illustrates count of points in the chromaticity plane as a two-dimensional histogram and shows Gaussian-like distributions. Each distribution corresponds to dye j is characterized by:

- mean value $(\alpha^{(j)}, \beta^{(j)})$ which determines the sampled spectral signature of the dye, and

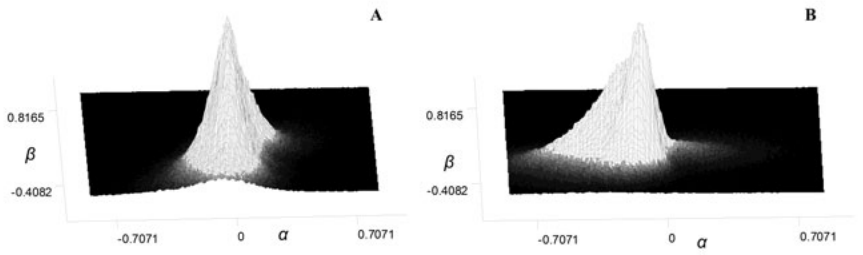


Figure 4.3: The result of transformation of image data shown in three-dimensional scatter plots Fig. 4.1C and Fig. 4.1D to two-dimensional chromaticity planes A and B, respectively. It can be observed from B that the eosin stained tissue is strongly affected by biochemical noise.

- covariance $\mathbf{Q}^{(j)}$ which is the measure of biochemical noise.

The Gaussian assumption [40] is reasonable from a statistical point of view as the noise term ϵ is random with zero-mean and there is no prior knowledge about variation of spectral signatures.

Gaussian assumption opens a possibility for maximum-likelihood estimation of the linear mixture model parameters $\mathbf{a}_j$ by means of expectation-maximizations. However, fitting n Gaussian distributions in a direct manner to data in the chromaticity space may not always match with the correct distributions. For instance, in the H/E example shown in Fig. 4.3B, the distribution associated with eosin is of low amplitude relative to the dominant distribution. In order to increase robustness to biochemical noise, Paper IV presents a method based on k-means clustering, a variation of expectation-maximization algorithm that assumes small, diagonal and equal covariances.

By taking the inverse of equation 4.2, under assumption $\|\mathbf{a}_j\|_1 = 1$, the mean values $(\alpha^{(j)}, \beta^{(j)})$ uniquely determine sampled spectral signatures of all dyes $\mathbf{a}_j$, $j = 1, ..., n$.

4.3 Piecewise linear decomposition

The previous section describes how to estimate model parameters in the chromaticity space. If the sampled spectral signatures of dyes form a well-conditioned mixing matrix $\mathbf{A} = [\mathbf{a}_1 \dots \mathbf{a}_n]^T$, relative dye concentrations are provided by linear decomposition, i.e., by using equation 3.18. Figure 4.4 shows the result of separation of hematoxylin and eosin stained histological section by linear decomposition.

On the other hand, spectral angle mapping [41] offers a stable solution when the mixing matrix is ill-conditioned. Unlike linear decomposition, angle mapping is a crisp pixel classification algorithm as it assigns each im-

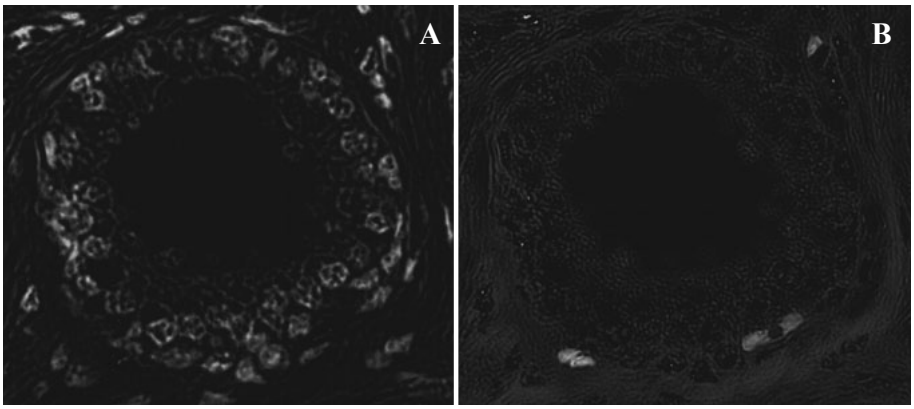


Figure 4.4: Estimated relative dye concentration maps of the hematoxylin (A) and eosin (B) stained tissue section shown in Fig. 4.1B.

aged position $\mathbf{p}$ to one of the n dyes with smallest angle

$$\theta_{\mathbf{p},j} = \arccos \frac{\mathbf{c_p} \cdot \mathbf{a}_j}{\|\mathbf{c_p}\|_2 \|\mathbf{a}_j\|_2}, \tag{4.4}$$

between the two sampled spectral signatures of the observed mixture $\mathbf{c_p}$ and sampled spectral signatures of dyes $\mathbf{a}_j$.

In addition to the two classical algorithms, piecewise linear decomposition is a novel method applicable to ill-conditioned cases, such as decomposition of the Gömöri trichromatic stain presented in Paper IV and detection of colocalization presented in Paper I. The first step is obtaining angular distances to sampled spectral signatures of dyes $\theta_{\mathbf{p},j}$ for the imaged position $\mathbf{p}$. Next, two dyes j_1 and j_2 are chosen such that

$$j_1, j_2, j_3 \in \{1,2,3\}, j_1 \neq j_2, j_2 \neq j_3, j_3 \neq j_1, \text{and } \theta_{\mathbf{p},j_1} < \theta_{\mathbf{p},j_3} \wedge \theta_{\mathbf{p},j_2} < \theta_{\mathbf{p},j_3}. \tag{4.5}$$

The estimated relative dye concentrations are

$$\begin{bmatrix} d_{\mathbf{p},j_1} \\ d_{\mathbf{p},j_2} \end{bmatrix} = \left(\left[\mathbf{a}_{j_1} \mathbf{a}_{j_2} \right]^T \left[\mathbf{a}_{j_1} \mathbf{a}_{j_2} \right] \right)^{-1} \left[\mathbf{a}_{j_1} \mathbf{a}_{j_2} \right]^T \mathbf{c_p}, \tag{4.6}$$

and

$$d_{\mathbf{p},j_3} = 0. \tag{4.7}$$

Detected projections of sampled spectral signatures $(\alpha^{(1)}, \beta^{(1)})$,$(\alpha^{(2)}, \beta^{(2)})$ and $(\alpha^{(3)}, \beta^{(3)})$ form a triangle in the chromaticity plane. The circumcenter of the triangle represents a sampled spectral signature where all dyes are mixed in equal proportions. Thus, if the triangle is acute, or ideally equilateral, the equation 3.18 holds as it is possible to unmix all three dyes. If the triangle is obtuse, the circumcenter lies outside of the triangle and rel-

ative dye concentrations are estimated by piecewise linear decomposition assuming that the two dyes associated acute vertices do not mix.

The following chapter presents a number applications where angle distributions α and chromaticity planes (α, β) were used in quantitative analysis of microscopy images. First, for extraction of sampled spectral signatures of dyes $\mathbf{a}_j$ and, eventually, to perform angle mapping, linear, non-linear or piecewise linear decomposition.

5. Results – applications in biotechnology and pathology

Data! Data! Data!
I can't make bricks without clay.
– Sherlock Holmes,
from "The Adventure of the Copper Beeches" by
Arthur Conan Doyle (1859-1930)

5.1 Quantification of colocalization (Paper I)

Colocalization occurs when two or more dyes are purposely mixed and appear physically co-located in the same imaged position. They combine spectrally to a new signature which is a linear combination of the signals from the two original dyes. For any number of spectral channels n_c, the mixing matrix $A = [\mathbf{a}_1\, \mathbf{a}_2\, \mathbf{a}_{1+2}]$ is non-invertible, thus neither model parameters nor relative densities can be obtained using least-squares based methods described in Chapter 3.

Paper I introduces angle histograms (equation 4.3) to estimate sampled spectral signatures of dyes $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_{1+2}$ in a dual-color image. The paper presents the basic example of colocalization ($n_c = 2$, $n = 3$) created using simulated images of an unknown source of fluorescent light. Therefore, it only deals with quantization noise and uniform background intensity offset. The simple method for parameter estimation is designed to extract dominant angles from predetermined angle intervals. Decomposition is implemented as angle mapping.

The result is a quantitative global measure of the degree of colocalization – coefficients M_1 and M_2. The colocalization coefficients represent approximation of the total amount of fluoresce of the colocalized dye $\mathbf{a}_{1+2}$ divided by the total amount of fluorescence by the pure dyes, determined by $\mathbf{a}_1$ and $\mathbf{a}_2$, respectively.

Conclusion: Angle mapping was compared to pixel classification methods based on intensity thresholding commonly used for quantification of colocalization. The method ensures that imaged positions containing either strong or weak signals from dyes belong to the same class. It outperforms related methods when applied to both simulated and real

microscopy images [42, 43, 44]. The study demonstrated in Paper III confirms this conclusion.

5.2 Suppression of cross-talk and background fluorescence (Paper I and Paper II)

Background fluorescence, also known as autofluorescence [45], and cross-talk or bleed-through [46, 47] are two problems in fluorescence microscopy often solved by spectral decomposition. When imaging biological specimens, the detected signal j in spectral channel k contains contributions from fluorescence originating from sources other than the imaged dye. This fluorescence could either come from the specimen itself (background fluorescence), or from dyes with partly overlapping emission spectra (cross-talk). According to equation 3.1, in order to resolve spectral components at least $n_c = 2$ distinct wavelength bandwidths have to be imaged **or** used for excitation.

In situations when spectral signatures are unknown, linear mixture model parameters are extracted from angle histograms and consequently used for compensating cross-talk, independent of the microscope hardware. Cross-talk shows up as shifts of the peaks in the histogram. Thus sampled spectra $\mathbf{a}_1$ and $\mathbf{a}_2$, represented by angles α_1 and α_2, respectively, can be estimated from the image data. In Paper I, this is implemented using the same histogram segmentation algorithm as for determining classification rules for angle mapping.

Paper II demonstrates that autofluorescence removal by carefully designed non-linear decomposition enables quantification of protein complexes in brain tissue detected by proximity ligation assay (see Section 2.4). In Chapter 3 fluorescence background was assumed to be equal to the black level offset. The paper addresses a situation when the assumption does not hold. While protein complexes are detected by probes excited at 563nm, lipofuscin and hemoglobin, main contributors to autofluorescence, have very broad spectral profiles. Background fluorescence was excited at 488nm providing an image free of point-source signals. Since images contain only a few signals per field of view, non-negative matrix factorization or direct segmentation of the angle histogram are inapplicable. Paper II presents an application-specific algorithm that provides parameters for decomposition. Finally, by applying non-linear decomposition the measure of relative dye concentration is lost. However, point-source signals are equally important irrespectively to their intensity and the level of background fluorescence, hence non-linear unmixing provides higher signals-to-noise ratio.

Conclusion: In addition to colocalization coefficients, the analysis of angle histograms presented in Paper I can be employed for estimation of

sampled spectral signatures in dual-channel fluorescence microscopy. By employing linear decomposition, cross-talk-free fluorescence microscopy images are generated. For the specific application presented in Paper II, the algorithm also saves time and space for storage of multispectral images as only two spectral channels need to be recorded and processed to remove autofluorescence. The non-linear decomposition algorithm proposed in Paper II for enhancement of point-source signals shows improvement in comparison to linear decomposition.

5.3 Classification of multicolored signals (Paper III)

Paper III presents a study of biomolecular methods for specific detection of single molecules and single molecule interactions. Specificity of the detection originates from the circular-DNA molecule created only if padlock or PLA probes bind to a well-defined DNA sequence or pair of proteins [19, 23]. Subnanometer-sized molecules are labeled and enlarged by means of rolling-circle amplification, hence increasing the local intensity and ultimately satisfying the Shannon-Nyquist sampling theorem. Further on, fluorescent colors are incorporated at a hybridization site and utilized for multiple detection of single molecules providing a unique spectral signature for each molecule or molecular interaction of interest. During rolling circle amplification, the characteristic spectral signature of a dye or a mixture of dyes is multiplied, and intensity ratios vary within boundaries determined by the level of biochemical noise.

Point-source signals are modeled by their shape as each signal, in addition to increased brightness, expands spatially. Paper III demonstrates a method that combines object based detection of point-source signals [48] and the global measure of colocalization presented in Paper I. The resulting angle histogram can be employed to classify multicolored signals or to estimate the level of biochemical noise.

Conclusion: Paper III demonstrates that the choice of the method for oligonucleotide synthesis does not affect ability of multiplexing. The paper presents a procedure for comparison of methods for quantification of protein-DNA interactions, which has also been employed by Weibrecht *et al.* [49].

5.4 Color decomposition of histological images (Paper IV)

Paper IV presents the method explained in the previous chapter applied to color decomposition of histopathological stains. The main challenge ad-

dressed in this paper is automated detection of sampled spectral signatures with a minimum of *a priori* information.

While the methods in Papers I-III initially collect data points in angle bins to form an angle histogram, this paper presents a method based on k-means clustering. The simple pattern analysis algorithm is employed to estimate sampled spectral signatures of dyes. In comparison to non-negative matrix factorization, which is a method implemented as an alternating least squares procedure [35], the proposed pattern analysis technique performs at least as well or better than non-negative matrix factorization. It can also detect collinear projections of sampled spectral signatures and thus address ill-conditioned cases. Naturally, very little or no stain in tissue makes all blind methods fail, which is a practical limitation of the approach.

In addition to decomposition of light-absorbing stains that follow the Beer-Lambert law, the paper offers a solution for treatment of light-scattering stains. Those stains, such as diaminobenzidine, do not adhere to the Beer-Lambert law and have to be removed in the pre-processing step. The paper also presents compensation for photon noise that particularly affects regions with low dye concentration.

Conclusion: Paper IV demonstrates that estimation of the linear model parameters by analyzing data in the chromaticity plane outperforms state-of-the-art algorithms both for well-conditioned cases (such as H/E) and ill-conditioned cases (Gömöri trichrome).

6. Discussion

I have had my results for a long time:
but I do not yet know how I am to arrive at them
– Carl Friedrich Gauss (1777-1855)

This chapter covers a wide range of topics ranging from description of limitations of the proposed linear model and method to a discussion on color spaces, sampling and that major advances will result from collaborative work.

6.1 Limitations of the linear mixture model

The transfer functions of optical imaging systems equations 3.3 and 3.9 have limitations. In particular they do not take into account the point spread function of the lens, which is also wavelength dependent [3], but assume lenses of infinite size. Thus, a complete model of the system would require integration over three spatial dimensions to compensate for different degrees of blurring when photons from imaged position $\mathbf{p}$ reach the detector.

In fluorescence microscopy, the main limitation of the model stems from colocated dyes that may interact. For example, Förster resonance energy transfer [50] introduces artifacts visible at high concentrations in the color cube [31]. In bright-field microscopy, the linear mixture model does provide quantitative information, i.e., relative dye concentrations, when light-absorbing dyes are used for specimen preparation. Nevertheless, it is common to use light-scattering stains which results in loss of quantitative information [12, 32, 51]. Paper IV proposes a solution to the problem, but also suggests that light-scattering stains (including diaminobenzidine) require a new model.

Epifluorescence and bright-field microscopes shown in Fig. 3.1 and 3.2, respectively, are wide-field microscopes. They use sensor array to acquire two-dimensional images, while confocal microscopes use photomultiplier tubes (single-pixel cameras) to acquire three-dimensional spatial information [3]. While wide-field microscopes suffer from both longitudinal and lateral chromatic aberrations, only longitudinal chromatic aberrations influence confocal microscope imaging. The methods for spectral and color decomposition presented in this thesis require images free of chromatic

aberrations, just as spectral unmixing or intensity-based quantification of colocalization algorithms. In wide-field microscopy, lateral chromatic aberrations may affect results significantly, particularly quantification of point-source signals. Algorithmic compensation for chromatic aberrations solves this problem [52, 53].

Finally, the definition of colocalization proposed in a previous chapter may be different from the description in Paper I: “The source of two different emission signals can often be physically located in the same area or *very near each other* in the final image due to their close proximity within the microscopic structure. This is known as colocalization." Although two single-colored objects laying next to each other may be considered as colocalized [54], an underlying assumption of the proposed method is existence of a significant number of dyes co-located in area smaller than area $\varsigma_{\mathbf{p}}$ projected to the specimen plane.

6.2 In light microscopy, there is no such thing as – hue

This thesis presents a method for decoupling light intensity from spectral information. It also considers tri-chromatic red-green-blue color sensors as one type of spectral imaging systems. In modern digital image processing, researchers tend to follow Goethe's description of color as a mixture of primaries as perceived by humans and refer to separation of color into hue and saturation as decoupling of spectral information. But in microscopy, the method based on the linear mixture model has been proven to be superior to hue-based solutions [55].

On the other hand, with the exception of a physically-based approach [56], computer vision researchers often prefer decoupling of intensity and spectral information by transforming input data to the L*a*b* color space. The transformation is modeled by the non-linear response of human eyes. Therefore, the chromaticity plane a*b* preserves color differences as perceived by humans. Since the most commonly observed objects in computer vision and multimedia are man-made (apart from the human skin color [57]) this transformation appears to be suitable. The question is: Should quantitative microscopy represent color as it is perceived by humans? And the answer supported by this thesis is – no. The chromaticity plane introduced in Chapter 4 preserves chromaticity differences as differences between sampled spectral signatures of dyes. It also opens a possibility of modeling and measuring biochemical noise as a finite sum of Gaussian distributions.

Finally, Newton left it unclear what he meant by “centers of gravity let Circles proportional to the number of Rays of each Colour in the given Mixture be describ'd” (see Chapter 2). But in the context of modern optical imaging, the following explanation is plausible: the color wheel is not a wheel (and

certainly should not be used to represent color as a hue-saturation couple); it is a projection of a seven-dimensional color cube to a two-dimensional space – simply for visualization purposes.

6.3 Uniform vs. irregular sampling

Hyperspectral cameras (Fig. 2.7A) can be used for identification of molecules with distinct line spectra. Uniform sampling of the light spectrum in such imaging systems is band limited by the diffraction grating and sampling occurs in accordance with the Shannon-Nyquist criterion (eq. 2.5). Thus, a sampled spectral signature with a large number of spectral channels for each imaged spatial position is recorded. The main drawback of hyperspectral imaging systems is poor signal-to-noise ratio caused by narrow band channels, i.e., high resolution of the system.

Several studies demonstrate that the number of spectral channels does not need to exceed the number of dyes to maintain separability of overlapping spectra and high signal-to-noise ratio [58, 59, 60]. Therefore, in fluorescence microscopy sampling of spectral information is characterized by assumed emission spectra of fluorescent dyes and it is in its nature – irregular. In addition, using a red-green-blue filters arranged into a form of a mosaic filter, the most common sensor in commercial cameras [33], is another example of irregular sampling of spectral information.

Thus, irregular sampling of spectral information prevails in all fields of science and technology. On the other hand, irregular sampling of temporal and spatial information [61] is still not very common in applied sciences or medicine where high resolution uniform sampling generates image data both for visualization and image analysis. Perhaps, the answer to why it is so common to sample the spatial and temporal domains uniformly lies in the fact that human perception affects the way engineers implement sensors. Apart from biologists' desire to study three-dimensional high resolution imagery, recent papers indicate that time-sequences [62] also play an important role in quantitative microscopy. And while large amount of data from the proliferation of sensors with ever increasing resolution is beneficial, it also represents one limiting factor for the 21st century computer architectures.

Guided by the following facts: (1) irregular sampling of light spectra did not show any major drawbacks; (2) information of interest is often distributed over more than two dimensions; (3) imaging time grows often linearly with the amount of data; (4) the amount of data grows exponentially with each new dimension; it appears reasonable to pose the following question: Could sparse sampling of temporal and spatial information [63] be a technique for efficient image acquisition that will revolutionize image processing that we know today?

6.4 Future of image processing in medicine and biology

The methods published in Paper III have already been employed for evaluation of molecular biology protocols, i.e., detection of individual sequence-specific protein-DNA interactions *in situ* [49]. Image processing in this work was not limited to analysis of visually acceptable images provided by biotechnologists (see Fig. 2.4), but it was an integral part of protocols for measuring the detection probe quality and accuracy (Fig. 6.1).

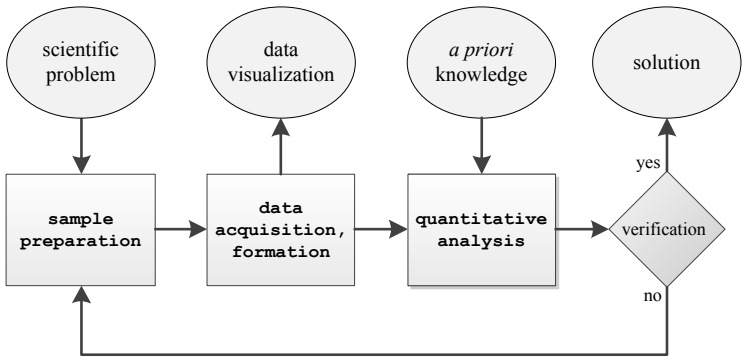


Figure 6.1: An example of collaborative work – information processing applied in all stages of research may minimize subjectivity in diagnosis and decrease bias.

Furthermore, as a consequence of accurate estimation of relative dye concentrations in Paper IV morphological structures are preserved. For example, Gleason grading of prostate cancer [64] uses a number of morphological features, from shape of glandular units to the size of nuclei. In addition, by using image analysis, computers can help pathologists by providing extensive statistics of a number of features, e.g., chromatin distribution in cell nuclei, an important feature for severity grading [65].

Section 2.4 describes two different philosophies in quantitative analysis, machine learning and analytical modeling. Today, the first approach prevails. And while it is cost effective, its conclusions are limited by the material provided for the training step. In medicine, this means that image analysis is limited by the performance of medical doctors [66] because statistical models based on experts' annotations and opinions are not always optimal. However, employing researchers from both biomedicine and engineering (see Fig. 6.1) to develop and adopt new stains and protocols jointly, derive mathematical models and test their limitations should be the approach.

Sammanfattning på svenska

Färgteori och mikroskopi har samexisterat i flera århundraden. I och med 1900-talets teknologiska framsteg har tidigare nära relaterade forskningsområden som färgteori, spektroskopi, mikroskopi och datoriserad bildanalys sammanfogats i sann synergi. Parallellt med detta har förståelsen för människans förmåga att tolka färginformation ökat. Detta har lett till såväl kvantitativa mått på färg som till modeller för hur vi uppfattar färg visuellt.

Med tanke på att målet med bildbehandling inom naturvetenskapliga tillämpningar och medicin är att säkerställa en objektiv analys, är denna avhandling fokuserad på färgteori inriktad på kvantitativ analys snarare än det subjektiva angreppssätt som dominerar forskningsområdet idag. Ett område där detta är särskilt angeläget är inom histopatologi, där målet är att undvika subjektiva bedömningar och öka tillförlitligheten för en diagnos och därmed uppnå bättre behandlingsplanering. Dessutom kan datoriserad bildbehandling dra nytta av matematiska modeller som kan bekräfta eller förkasta vetenskapliga hypoteser på ett tidigt forskningsstadium genom att öka noggrannheten och analyskapaciteten och minska subjektiviteten.

Avhandlingen presenterar en modell för spektral bildanalys som kan tillämpas såväl på fluorescens- som ljusfältsmikroskopi. Genom att tillämpa den inversa modellen erhålls uppskattningar av den relativa koncentrationen för var och en av komponenterna i den observerade blandningen av infärgningar. Modellens parameterskattning grundar sig på att vi skiljer på intensitetsinformation och spektral information. Denna nya spektrala uppdelningsmetod är uppbyggd i tre steg: 1) Modellering av foton- och halvledarbrus leder till parametrar för utjämning av signalerna 2) Transformering av bilden till ett kromaticitetsplan tar bort intensitetsvariationer samtidigt som skillnader i färgnyans bevaras. 3) styckvis linjär uppdelning kombinerar fördelarna av spektral mappning och linjär uppdelning och leder till relativa koncentrationer av färgämnen.

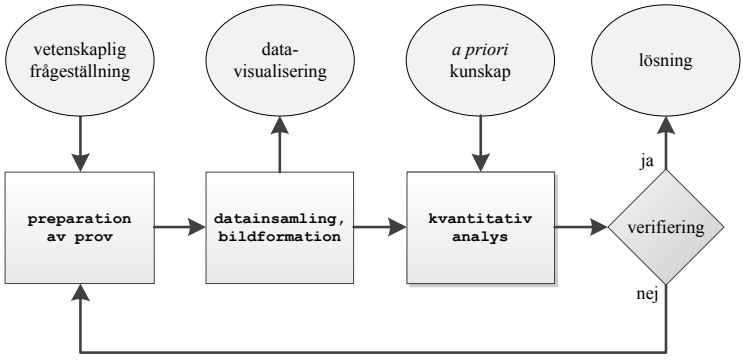
Metoderna inkluderar en matematisk modell och algoritmer för automatiserad parameterskattning. Ett antal ljusmikroskopitillämpningar viktiga för kvantitativ mikroskopi presenteras:

- Kvantifiering av närliggande fluorescenssignaler. När flerfärgade prover avbildas händer det ofta att fluorescenssignaler från olikfärgade men närliggande mikroskopiska strukturer hamnar i samma fysiska bildpunkt. Detta leder till ett spektralt uppdelningsproblem som

inte kan lösas med klassisk egenvektorbaserad multivariat analys. I avhandlingen presenteras ett spektralt vinkelhistogram; ett nytt verktyg för visualisering och kvantitativ analys inom mikroskopi. Vinkelhistogrammet separerar spektral information från intensitetsinformation och tar hand om de illa-konditionerade uppdelningsproblem som nämns ovan.

- Spektral uppdelning vid autofluorescens och signalöverhörning. Bakgrundsfluorescens, också kallad autofluorescens, och signalöverhörning är två problem inom fluorescensmikroskopi som har sitt ursprung i liknande fysiska fenomen, och ofta löses med hjälp av spektral uppdelning. När biologiska prover avbildas innehåller den detekterade signalen ofta bidrag från andra källor än den avbildade fluorokromen. Dessa signalbidrag kommer kommer antingen från provet självt (bakgrundsfluorescens) eller från fluorokromer med delvis överlappande emissionsspektra (signalöverhörning). För att kunna separera signaler från olika källor krävs avbildning av åtminstone två olika våglängdsintervall, alternativt avbildning med två olika excitationsspektra. För klassisk spektraluppdelning krävs spektrala signaturer av individuella komponenter, och dessa spektrala signaturer måste avbildas innan spektraluppdelning är möjlig. I situationer där spektrala signatur är okända kan vinkelhistogram användas för att bestämma parametrar för både linjär och icke-linjär spektraluppdelning. En annan fördel med metoden är att enbart ett fåtal spektrala kanaler behöver avbildas och behandlas – något som spar både tid och lagringsutrymme.
- Färgbaserad vävnadsklassificering i ljusmikroskopi. Genom att tillämpa den inversa modellen som extraherar spektrala komponenter från ett vävnadsprov kan en färgkoncentrationskarta skapas. Variationer i färgintensitet är därmed borträknade, och artefakter från provpreparation och infärgning reduceras. Dessutom kan modellen kombineras med statistisk modellering av brus och mjuka klassificeringsmetoder möjliggör linjär färgseparation.

Slutligen presenteras metoder för kvantitativ bildanalys tillämpade inom ett samarbetsprojekt med biomedicinska forskare. Traditionellt sett brukar forskare inom bildanalys bli inblandade i samarbetsprojekt först efter att de biomedicinska metoderna utvecklats, och ibland även efter att bilddata insamlats. Detta gör att resultaten begränsas av att bilderna inte är anpassade för bildanalys. För att övervinna begränsningarna med detta angreppssätt har vi utvecklat metoder för bildbehandling parallellt med metoder för att kvantifiera interaktioner mellan enskilda molekyler visualiserade med hjälp av flerfärgade signaler. Vi beskriver hur bildbehandling används för att utvärdera biokemiska metoder, inklusive provpreparering, avbildning och kvantitativ analys. På liknande sätt har överföringen av färginformationen i bilderna till ett kromaticitetsplan



Figur 1: Informationsbehandling inom medicin och biologi.

möjliggjort en teoretisk beskrivning av en perfekt infärgning. Dess referensfärger skall vara belägna i hörnen på kromaticitetsplanet och alla vävnadstyper skall ha lika stort upptag av färg (ingen övermättnad, inga bleka infärgningar). Dessa kriterier hjälper patologer att utveckla nya infärgningar som är optimerade primärt för kvantiativ analys med datorer och i andra hand för att fungera för det mänskliga synsinnet. Figur 1 illustrerar flödet i den bildbehandlingsprocess som skapades för att uppfylla behoven inom detta interdisciplinära projekt.

Acknowledgements

"You have enemies? Good. That means you've stood up for something, sometime in your life," once said Winston Churchill. But this thesis would have never been written if there were not my friends – my family, my colleagues and my teachers whom I would like to acknowledge at the end of this journey.

In order of appearance:

- My parents, Nada and Radojko, for giving me all the freedom I ever wanted and their unconditional support, for sharing their great knowledge and teaching me how to seek new knowledge.

- My beloved Jasna for unconditional love, being by my side when it rains and pours and supporting me in all ways imaginable.

- Teachers at the Faculty of Electrical Engineering in Belgrade, Dragi Dujković, Branimir Reljin and Srbijanka Turajlić, who initiated my interests in control theory and image processing, and encouraged my early research efforts.

- **Most of all**, my doctoral advisors: Carolina Wählby for inviting me to work on my first real research project, a warm welcome to Uppsala, helping me to make a smooth turn from electrical engineering to engineering in biology and learn the language of biotechnology; Ewert Bengtsson for introducing me to medical engineering and for opening the door to the "fifth" dimension of image processing; Ingrid Carlbom for recognizing my skills and introducing me to a prostate cancer grading problem, for being persistent in teaching me the significance of liberal arts and sharing a number of techniques for scientific writing, I am happy to announce: I do not hate writing anymore!

- Joakim Lindblad, Gunilla Borgefors and Cris Luengo, who were never formally associated with my studies, for playing an important role in my graduate education. Christer Busch for teaching me Gleason grading – I

almost feel confident to diagnose prostate cancer.

- My biotech project co-workers and fellow PhD students, Irene Weibrecht and Tim Conze, for our amazing collaboration as well as their advisors Ola Söderberg and Ulf Landegren for supporting our joint work. Amin Allalou for sharing his magnificent algorithm for detection of point-source signals.

- My officemates, Amalka Pinidiyaarachchi, Andreas Kårsnäs and Jimmy Azar, for their companionship. J. C.Azar is also acknowledged for his help with pattern analysis techniques in Paper IV.

- Hamid Sarve, M. Khalid K. Niazi, Vladimir Ćurić and Pontus Olsson for being more than colleagues and spending hours and hours discussing everything about anything.

- Lena Wadelius and Olle Eriksson for making me feel like being at home even when I am at work. Bo Nordin, my late lunch buddy (Rullan, 13:55).

- All grad students and junior faculties from the corridor for contributing to this exciting learning environment (a.k.a. CBA): Kristin Norell, Erik Vidholm, Maria Axelsson, Patrick Karlsson Edlund (also for teaching me his teaching philosophy), Filip Malmberg, Stina Svensson, Robin Strand, Nataša Sladoje, Magnus Gedda, Bettina Selig (also for keeping us supplied with finest German chocolates during these long, cold winters), Erik Wernersson (also for being a true image processor), Patrik Malm (also for lessons in cross-country skiing), Gustaf Kylberg, Ida-Maria Sintorn, Lennart Svensson, Martin Ericsson, Anders Brun, Hyun-Ju Choi, Elisabeth Linnér (also for still believing in bright future for academia in Sweden) and Azadeh Fakhrzadeh, as well as Jens Hedrich for sharing my enthusiasm for color theory.

- My fellow grad students from the IT dept: Carl Nettelblad, Mikael Laaksoharju and my great Vemdalen roommates, Soma Tayamon, Daniel Elfverson, Josefin Ahlkrona, Marcus Björk and Per Mattsson. Thank you for all the *fun* we had together, one way or another.

- All the undergraduate students I did not bore to death with mathematical morphology and Fourier analysis! Some of them later became dear colleagues of mine after they completed the course project: Ulf Hammarqvist, Fredrik Wahlberg and the Nysjö brothers. Viktor Smedby, make sure not to lose your enthusiasm!

A special additional acknowledgement is granted to Cris Luengo, Vladimir Ćurić, Jasna Pruner, Ewert Bengtsson, Carolina Wählby and Ingrid Carlbom who contributed to this thesis with proof reading and providing a number of useful suggestions. I am grateful to Ulrika A. Larsson from the Uppsala University Hospital who helped with tissue preparation (Chapter 4). Three-dimensional data in Paper IV and Chapter 4 was visualized by using *3D Color Inspector*. The staff at *Publicering och grafisk service* is acknowledged for understanding that the deadlines in academia are of relative importance.

Another special note to all colleagues who often kept me company during late nights at work, in particular Amin, Gunilla, Khalid and Ingrid. I have a dream... that one day we will end the tyranny of morning people! 10am is time to be at home (and sleep), 10pm is the right time to work on new ideas.

In Uppsala, 25 October 2011, 4:34am

Milan

Bibliography

[1] W. K. Pratt, *Digital Image Processing: PIKS Inside*. New York, NY, USA: Wiley-Interscience, 3rd ed., 2001.

[2] J. Bresenham, "Algorithm for computer control of a digital plotter," *IBM Systems Journal*, vol. 4, no. 1, pp. 25–30, 1965.

[3] K. Carlsson, *Light Microscopy*. Stockholm, Sweden: Department of Applied Physics, Royal Institute of Technology, 2007.

[4] I. Newton, *Opticks: Or a Treatise of the Reflections, Refractions, Inflections & Colours of Light*. 4th ed., 1730.

[5] F. Burwick, *The damnation of Newton : Goethe's color theory and romantic perception*. Berlin, Germany: Walter de Gruyter, 1st ed., 1986.

[6] J. P. Frisby and J. V. Stone, *Seeing: The Computational Approach to Biological Vision*. MIT Press, 2nd ed., 2010.

[7] J. K. Bowmaker and H. J. Dartnall, "Visual pigments of rods and cones in a human retina," *The Journal of Physiology*, vol. 298, pp. 501–511, 1980.

[8] N. Riebe and F. Steinle, "Exploratory experimentation: Goethe, land, and color theory," *Physics Today*, vol. 55, no. 7, pp. 43–49, 2002.

[9] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 3rd ed., 2008.

[10] J. Marshall and J. Oberwinkler, "Ultraviolet vision: The colourful world of the mantis shrimp," *Nature*, vol. 401, pp. 873–874, 1999.

[11] G. Chartier, *Introduction to Optics*. New York, NY, USA: Springer Science+Business Media, Inc., 1st ed., 2005.

[12] Y. Garini, I. T. Young, and G. McNamara, "Spectral imaging : Principles and applications," *Cytometry*, vol. 747, no. 8, pp. 735–747, 2006.

[13] J. C. Mullikin, L. J. van Vliet, H. Netten, F. R. Boddeke, G. van der Feltz, and I. T. Young, "Methods for ccd camera characterization," *SPIE Proceedings on Image Acquisition and Scientific Imaging Systems*, vol. 2173, pp. 73–84, 1994.

[14] G. Polder and G. W. van der Heijden, "Calibration and characterization of spectral imaging systems," *SPIE Proceedings on Multispectral and Hyperspectral Image Acquistition I*, vol. 2448, no. 10, 2001.

[15] I. T. Young, J. J. Gerbrands, and L. J. van Vliet, *Fundamentals of Image Processing*. Delft, Netherlands: Delft University of Technology, 2nd ed., 2007.

[16] "Noise analysis in operational amplifier circuits," *Texas Instruments – Application Report SLVA043B*.

[17] K. Carlsson, *Imaging Physics*. Stockholm, Sweden: Department of Applied Physics, Royal Institute of Technology, 2009.

[18] I. B. Buchwalow and W. Böcker, *Immunohistochemistry*. Springer-Verlag.

[19] I. Weibrecht, K.-J. Leuchowius, C.-M. Clausson, T. Conze, M. Jarvius, M. Howell, M. Kamali-Moghaddam, and O. Söderberg, "Proximity ligation assays: a recent addition to the proteomics toolbox," *Expert Review of Proteomics*, vol. 7, no. 3, pp. 401–409, 2010.

[20] T. K. Kerppola, "Design and implementation of bimolecular fluorescence complementation (BiFC) assays for the visualization of protein interactions in living cells," *Nature Protocols*, vol. 1, no. 4, pp. 1278–1286, 2006.

[21] G. Bussolati and E. Leonardo, "Technical pitfalls potentially affecting diagnoses in immunohistochemistry," *Journal of Clinical Pathology*, vol. 61, no. 11, pp. 1184–1192, 2008.

[22] J. A. Ramos-Vara, "Technical aspects of immunohistochemistry," *Veterinary Pathology*, vol. 42, pp. 405–426, 2005.

[23] M. Nilsson, F. Dahl, C. Larsson, M. Gullberg, and J. Stenberg, "Analyzing genes using closing and replicating circles," *Trends in Biotechnology*, vol. 24, no. 2, pp. 83–88, 2006.

[24] Q. Wu, F. Merchant, and K. Castleman, *Microscope Image Processing*. Academic Press, 2008.

[25] K. R. Castleman, "Concepts in imaging and microscopy: Color image processing for microscopy," *The Biological Bulletin*, vol. 194, no. 2, pp. 100–107, 1998.

[26] H. Choi, K. R. Castleman, and A. C. Bovik, "Color compensation of multicolor fish images," *IEEE Transactions on Medical Imaging*, vol. 28, no. 1, pp. 129–136, 2009.

[27] A. C. Ruifrok and D. A. Johnston, "Quantification of histochemical staining by color deconvolution," *Analytical and quantitative cytology and histology*, vol. 23, no. 4, pp. 291–299, 2001.

[28] T. Zimmermann, "Spectral imaging and linear unmixing in light microscopy," in *Advances in Biochemical Engineering/Biotechnology, Vol. 95* (J. Rietdorf, ed.), pp. 245–265, Berlin Heidelberg, Germany: Springer-Verlag, 2005.

[29] T. Pengo, A. M. noz Barrutía, and C. O. de Solórzano, "Spectral unmixing of multiply stained fluorescence samples," in *Microscopy: Science, Technology, Applications and Education* (A. Mendez-Vilas and J.Diaz, eds.), pp. 2079–2087, Badajoz, Spain: Formatex Research Center, 2010.

[30] Y. Garini, A. Gil, I. Bar-Am, D. Cabib, and N. Katzir, "Signal to noise analysis of multiple color fluorescence imaging microscopy," *Cytometry*, vol. 35, no. 3, pp. 214–226, 1999.

[31] K. R. Castleman, R. Eils, L. Morrison, J. Piper, K. Saracoglu, M. A. Schulze, and M. R. Speicher, "Classification accuracy in multiple color fluorescence imaging microscopy," *Cytometry*, vol. 41, no. 2, pp. 139–147, 2000.

[32] W. W. Parson, *Modern Optical Spectroscopy*. Berlin Heidelberg, Germany: Springer-Verlag, 2nd ed., 2007.

[33] A. Koschan and M. Abidi, *Digital Color Image Processing*. Hoboken, NJ, USA: John Wiley & Sons, Inc., 1st ed., 2008.

[34] M. E. Dickinson, G. Bearman, S. Tille, R. Lansford, and S. E. Fraser, "Multi-spectral imaging and linear unmixing add a whole new dimension to laser scanning fluorescence microscopy.," *Biotechniques*, vol. 31, no. 6, 2001.

[35] L. Eldén, *Matrix Methods in Data Mining and Pattern Recognition*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2007.

[36] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Advances in Neural Information Processing Systems 13* (T. K. Leen, T. G. Dietterich, and V. Tresp, eds.), pp. 556–562, MIT Press, 2001.

[37] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 44–57, 2002.

[38] J. R. Mansfield, C. Hoyt, and R. M. Levenson, "Visualization of microscopy-based spectral imaging data from multi-label tissue sections," *Current Protocols in Molecular Biology*, pp. 14.19.1–14.19.15, 2008.

[39] J. Foley, A. van Dam, S. Feinera, and J. Hughes, *Computer Graphics: Principles and Practice*. Reading, MA, USA: Addison-Wesley, 2nd ed., 1992.

[40] P. Stoica and P. Babu, "The Gaussian data assumption leads to the largest Cramér-Rao bound," *IEEE Signal Processing Magazine*, vol. 28, no. 3, pp. 132 – 133, 2011.

[41] F. A. Kruse, A. B. Lefkoff, J. B. Boardman, K. B. Heidebrecht, A. T. Shapiro, P. J.Barloon, and A. F. H.Goetz, "The spectral image processing system (SIPS) – interactive visualization and analysis of imaging spectrometer data," *Remote Sensing of Environment*, vol. 44, pp. 145–163, 1993.

[42] E. M. M. Manders, J. Stap, G. J. Brakenhoff, R. van Driel, and J. A. Aten, "Dynamics of three-dimensional replication patterns during the s-phase, analysed by double labelling of DNA and confocal microscopy," *Journal of Cell Science*, vol. 103, no. 3, pp. 857–862, 1992.

[43] E. M. M. Manders, F. J. Verbeek, and J. A. Aten, "Measurement of co-localization of objects in dual-color confocal images," *Journal of Microscopy*, vol. 169, no. 3, pp. 375–382, 1993.

[44] S. V. Costes, D. Daelemans, E. H. Cho, Z. Dobbin, G. Pavlakis, and S. Lockett, "Automatic and quantitative measurement of protein-protein colocalization in live cells," *Biophysical Journal*, vol. 86, pp. 3993–4003, 2004.

[45] F. Woolfe, M. Gerdes, M. Bello, X. Tao, and A. Can, "Autofluorescence removal by non-negative matrix factorization," *IEEE Transactions on Image Processing*, vol. 20, no. 4, pp. 1085–1093, 2011.

[46] K. Carlsson and K. Mossberg, "Reduction of cross-talk between fluorescent labels in scanning laser microscopy," *Journal of Microscopy*, vol. 167, no. 1, pp. 23–37, 1992.

[47] C. L. L. Hendriks, S. V. E. Keränen, M. D. Biggin, and D. W. Knowles, "Automatic channel unmixing for high-throughput quantitative analysis of fluorescence images," *Optics Express*, vol. 15, no. 19, pp. 12306–12317, 2007.

[48] A. Allalou, A. Pinidiyaarachchi, and C. Wählby, "Robust signal detection in 3D fluorescence microscopy," *Cytometry Part A*, vol. 77A, no. 1, pp. 86–96, 2009.

[49] I. Weibrecht, M. Gavrilovic, L. Lindbom, U. Landegren, C. Wählby, and O. Söderberg, "Visualising individual sequence-specific protein-DNA interactions in situ," *New Biotechnology*, 2011.

[50] F. G. E. Cremazy, E. M. M. Manders, P. I. H. Bastiaens, G. Kramera, G. L. Hager, E. B. van Munster, P. J. Verschure, T. W. J. G. Jr., and R. van Driel, "Imaging in situ protein-DNA interactions in the cell nucleus using FRET-FLIM," *Experimental cell research*, vol. 309, no. 2, pp. 390–396, 2005.

[51] C. M. van der Loos, "Multiple immunoenzyme staining: Methods and visualizations for the observation with spectral imaging," *Journal of Histochemistry & Cytochemistry*, vol. 56, no. 4, pp. 313–328, 2008.

[52] E. M. M. Manders, "Chromatic shift in multicolour confocal microscopy," *Journal of Microscopy*, vol. 185, no. 3, pp. 321–328, 1997.

[53] M. Kozubek and P. Matula, "An efficient algorithm for measurement and correction of chromatic aberrations in fluorescence microscopy," *Journal of Microscopy*, vol. 200, no. 3, pp. 206–217, 2000.

[54] J. S. S. V. Hoeve and J. G. G. Borst, "Delayed appearance of the scaffolding proteins PSD-95 and homer-1 at the developing rat calyx of held synapse," *The Journal of Comparative Neurology*, vol. 518, no. 22, pp. 4581–4590, 2010.

[55] A. C. Ruifrok, R. L. Katz, and D. A. Johnston, "Comparison of quantification of histochemical staining by hue-saturation-intensity (HSI) transformation and color-deconvolution," *Applied Immunohistochemistry and Molecular Morphology*, vol. 11, no. 1, pp. 85–91, 2003.

[56] G. Klinker, S. Shafer, and T. Kanade, "A physical approach to color image understanding," *International Journal of Computer Vision*, vol. 4, no. 1, pp. 7–38, 1991.

[57] G. Pingali, G. Tunali, and I. Carlbom, "Audio-visual tracking for natural interactivity," in *Proceedings of the seventh ACM international conference on Multimedia (Part 1)*, MULTIMEDIA '99, (New York, NY, USA), pp. 373–382, ACM, 1999.

[58] R. Neher and E. Neher, "Optimizing imaging parameters for the separation of multiple labels in a fluorescence image," *Journal of Microscopy*, vol. 213, no. 1, pp. 46–62, 2004.

[59] E. Barash, S. Dinn, C. Sevinsky, and F. Ginty, "Multiplexed analysis of proteins in tissue using multispectral fluorescence imaging," *IEEE Transactions on Medical Imaging*, vol. 29, no. 8, pp. 1457–1462, 2010.

[60] L. E. Boucheron, Z. Bi, N. R. Harvey, B. S. Manjunath, and D. L. Rimm, "Utility of multispectral imaging for nuclear classification of routine clinical histopathology imagery," *BMC Cell Biology*, vol. 8, 2007.

[61] J. Edwards, "Focus on compressive sensing," *IEEE Signal Processing Magazine*, vol. 28, no. 2, pp. 11–13, 2011.

[62] Y. Hu, J. Carmona, and R. Murphy, "Application of temporal texture features to automated analysis of protein subcellular locations in time series fluorescence microscope images," *Proceedings of IEEE International Symposium on Biomedical Imaging*, pp. 1028–1031, April 2006.

[63] T. Blu, P. L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot, "Sparse sampling of signal innovations," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 31–40, 2008.

[64] M. B. Amin, D. J. Grignon, P. A. Humphrey, and J. R. Srigley, *Gleason Grading of Prostate Cancer*. Philadelphia, PA, USA: Lippincott Williams and Wilkins, 1st ed., 2004.

[65] B. Nielsen, F. Albregtsen, and H. E. Danielsen, "Statistical nuclear texture analysis in cancer research: a review of methods and applications," *Critical Reviews in Oncogenesis*, vol. 14, no. 2-3, pp. 89–164, 2008.

[66] J. Epstein, W. A. Jr, M. Amin, and L. Egevad, "Update on the Gleason grading system for prostate cancer: Results of an international consensus conference of urologic pathologists," *Advances in Anatomic Pathology*, vol. 13, no. 1, pp. 57–59, 2006.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 876

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology.

Distribution: publications.uu.se
urn:nbn:se:uu:diva-160574

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2011