Chapter 5
Short-Wavelength Infrared Hyperspectral Imaging for Biomedical Applications

Lise Lyngsnes Randeberg
Norwegian University of Science and Technology (NTNU), Trondheim, Norway

Julio Hernández
HySpex, Norsk Elektro Optikk AS, Oslo, Norway

Emilio Catelli
University of Bologna, Ravenna Campus, Italy and Norwegian University of Science and Technology (NTNU), Trondheim, Norway

5.1 Introduction

Hyperspectral imaging (HSI) is an optical technique combining high spectral and spatial resolutions into one modality. HSI is noninvasive, label-free (does not require dyes), and flexible and adaptable—which allows for imaging across a wide wavelength range.[1] The data are collected in the form of a hypercube consisting of one spectral and two spatial dimensions (Fig. 5.1). Occasionally, time or depth might be a fourth axis, adding to the multidimensionality of the data. Technical realization of a hyperspectral system and the data collection approach will depend on the application, as will the choice of tools for data analysis. A hyperspectral datacube can be collected using a variety of imaging geometries and systems, ranging from simple, as in using band-pass filters to image a portion of the spectrum, to more sophisticated, as in using Fourier imagers and imaging spectrometers.[1–14] The technology, which was originally developed for remote sensing, is now being adopted and adapted into fields that require spatially resolved spectroscopic information. This idea of combining high spectral resolution with imaging can be further combined with almost any other imaging technique and measurement geometry, including wide-field imaging of larger objects and endoscopic imaging.[1,15–20] As a result, HSI has recently been explored for a variety of medical applications and is slowly becoming a part of the standard biomedical toolbox.

By employing adequate optical elements and detectors, HSI can be used with light at wavelengths in the ultraviolet (UV) to thermal infrared range.[1,13,21,22] The short-wavelength infrared (SWIR) spectral range, covering ~900–2500 nm (Fig. 5.2), has been used in research and industrial fields such as geology and mining, materials recycling, food quality monitoring, solar cell inspection, and thin-film deposition.[13,23–25] This spectral range has

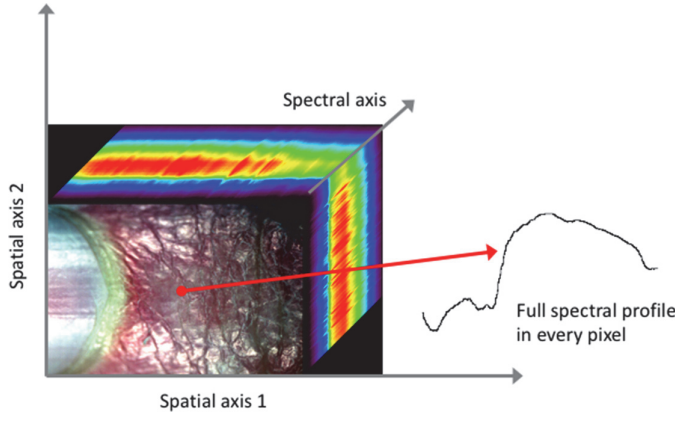

Figure 5.1 A hyperspectral datacube consists of at least two spatial dimensions and one spectral dimension. In some cases, time or depth might be added as a fourth coordinate, adding to the multidimensionality of the data.

Figure 5.2 Electromagnetic spectrum, from left to right: VIS (380–780 nm), NIR (780–900 nm), SWIR (900–2500 nm), mid-wave infrared (MIR: 2500–5000 nm), and long-wave infrared (LWIR: 8000–12,000 nm). The wavelength range 5000–8000 nm is defined as a stop band due to very high absorption and is therefore omitted.

also been utilized in biomedical applications and has several advantages. Typically, the interactions of SWIR and biological tissue result in lower autofluorescence, less scattering of light, and longer penetration depth compared to visible (VIS) and near-infrared (NIR) light ranges. The absorption of hemoglobin is far less in the SWIR than in the VIS/NIR part of the spectrum.

All hyperspectral information has a similar structure, which consists of spatially distributed spectral information. The data from different sources can be analyzed using the same or similar analytic tools,[13,23–25] regardless of the application. In addition, the imaging technologies are typically flexible and adaptable. Although there are hyperspectral instruments for specific clinical and research applications, most scientific-grade instruments from the major vendors can be used across a variety of disciplines, including biomedical research. Currently, HSI is more frequently incorporated as an additional feature to other imaging techniques—such as spatial frequency domain imaging (SFDI), coherent anti-Stokes Raman scattering (CARS) microscopy, and optical coherence tomography (OCT).[15,26–33]

Despite the trend of combining HSI with other imaging modalities, the available literature has not focused on SWIR optical imaging as such. Instead, HSI is associated with medical applications in general[1] or specific applications[5,21] such as cancer diagnostics. Several authors have explored the available SWIR hyperspectral hardware, including available detectors and other experimental details.[1,9,10,13,21,25,34] On the analysis side, the complexity and variety of the applied methods have increased.[35–55]

In this chapter, we give a brief overview of the technology and the field of HSI. We highlight the use of HSI in the SWIR spectral range (900–2500 nm) for biomedical applications. We focus on HSI data collection, acquisition, and processing, with an emphasis on chemometric and deep-learning methods. Some general guidelines and suggestions for selecting an adequate HSI imaging setup are also provided, together with some recommendations for improving the data acquisition process.

5.2 Medical Hyperspectral Imaging

HSI techniques have shown clinical potential in characterizing and detecting near-surface structures and anomalies.[22,56–58] The number of published papers on HSI has increased substantially recently as the technology has advanced. Currently, approximately two-thirds of the papers in the PubMed database mentioning the term hyperspectral or HSI were published after 2018 despite the fact that the first medical articles on HSI were published as early as the late 1990s and early 2000s.[59–72]

Traditionally, the VIS (380–780 nm) and NIR (780–1100 nm) spectral ranges were used for medical HSI, even though an increasing number of examples of SWIR (above 1100 nm) applications had been published.[73–83] Recently, SWIR cameras and detectors have become more available, and our understanding of tissue optical properties in this range has also increased.[56,84–86] Figure 5.3 shows some examples of SWIR-HSI for biomedical applications.

The published work in the literature can be classified and grouped according to various principles. It can be grouped according to imaging technology, i.e., microscopy, endoscopy, or wide-field imaging. It can also be grouped according to the purpose of the measurement: whether it aims at tissue characterization, diagnosis, surgical guidance, or treatment monitoring. Alternatively, the literature can also be grouped according to the physiology, anatomy, or biological system being studied. The latter presentation is the system used in Fig. 5.3.

Due to the optical properties of biological tissues, HSI is limited to surfaces, near-surface structures, or samples/tissues of penetrable thicknesses.

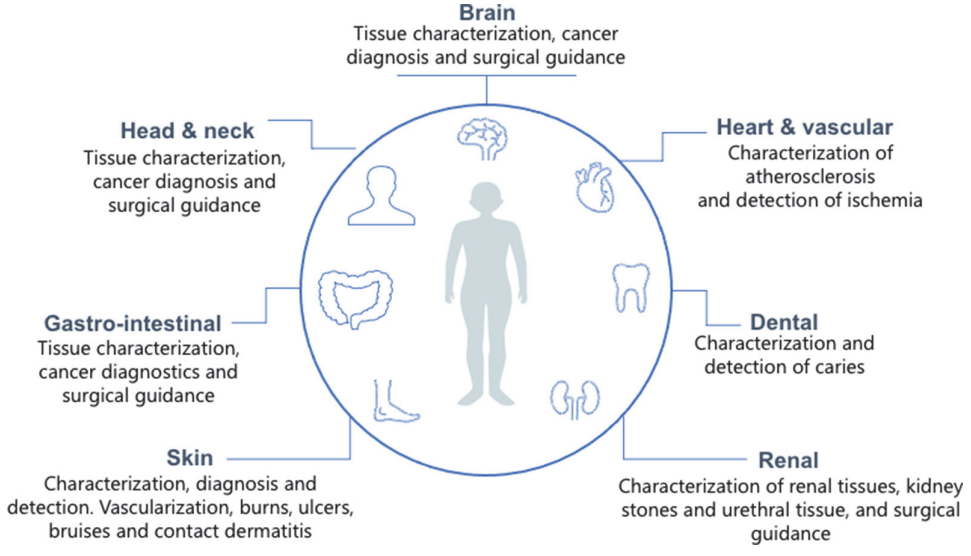

Figure 5.3 Examples of biomedical applications using SWIR-HSI.[1,5,21,38,48,81,87–95]

Thus, HSI is not a direct competitor of other medical imaging modalities such as ultrasound, magnetic resonance imaging, or X-rays; it is a complementary or additional technique.

Lu and Fei[1] give a good overview of the first 10–15 years of medical HSI. In their review, they discuss eight examples of HSI in the SWIR to MWIR spectral ranges from 1000 nm up to 5500 nm.[73,77–80,96–99] In one of their examples, they show how HSI can be utilized in the detection of intestinal ischemia and gastric cancer.[78,79] They also give examples of the characterization of dental caries, characterization of pharmaceuticals, and studies on breast cancer and cervical pathology in the MWIR range.[80,98,99]

Application of HSI to cancer diagnostics has grown substantially since 2014, as shown by Halicek et al.[21] In their article, they consider the spectral range from 390–2500 nm and discuss *ex vivo* classification of colon cancer,[79] *in vivo* classification of brain cancer,[100] and *ex vivo* classification and delineation of gastric cancer within the SWIR. Similarly, Baltussen et al.[81] show that a combination of VIS–NIR and SWIR imaging provides better results from classifying fatty tissue, colorectal wall tissue, and tumor tissue compared to imaging each wavelength region separately, when measured through a laparoscope (Fig. 5.4).

Other authors, such as Gao and Smith,[9] provide an overview of HSI in microscopy, with an emphasis on technology. In 2020, Ortega et al. presented a comprehensive review of digital and computational pathology.[5] In their review, they present "methods and uses of HSI/MSI for staining and color correction, immunohistochemistry, autofluorescence, and histopathological diagnostic research. Studies include hematology, breast cancer, head and neck cancer, skin cancer, and diseases of central nervous, gastrointestinal, and

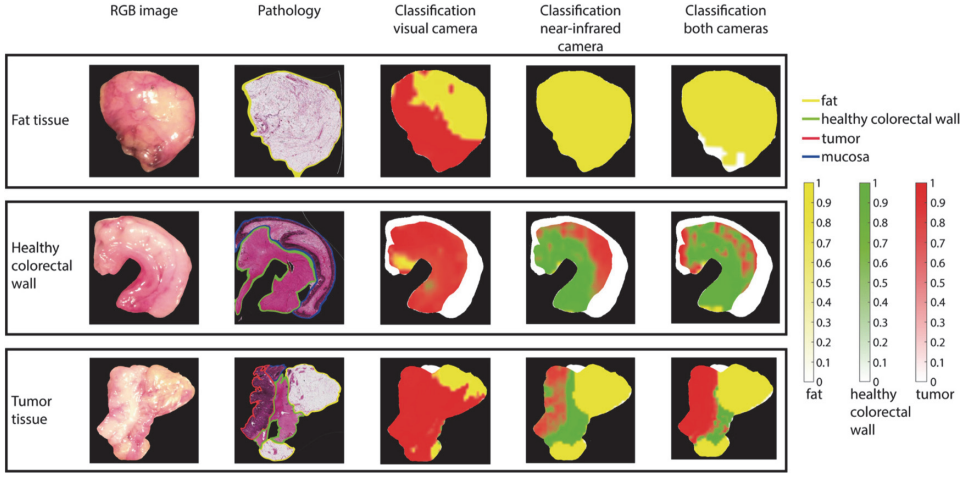

Figure 5.4 Identification of fatty tissue, colorectal tissue, and tumor tissue using HSI (RGB: red–green–blue). (Reprinted from Baltussen et al.[81])

genitourinary systems." However, despite their extensive review, Ortega et al.[5] and Gao and Smith[9] list few papers on SWIR, leaving SWIR as a relativity unexplored area of HSI and an area for future efforts.

Gowen et al.[19] review applications of HSI in microbiology and cover several applications in the SWIR range, but the applications are mainly related to food science. However, applications such as detection of fungi on surfaces and pathogenic bacteria (such as *Staphylococcus epidermidis* and *Enterobacteria* in water, and biofilms from *Escherichia coli* on microfluidic cells) are highly relevant for biomedical applications.[101–105] In general, food sciences have provided a large number of relevant studies on characterization and detection of bacteria that could be adapted to a medical context.

Skin—the most accessible organ on the human body and one of the largest organs—is the organ most explored by HSI, including characterization of healthy tissue and conditions of pathologies affecting the skin. Initial studies include examples on circulatory changes and skin oxygenation in diabetic feet.[93] Allergic contact dermatitis, burns, and ulcers are more recent examples of HSI of skin[75,93,94] (Fig. 5.5). Randeberg[75] presents hyperspectral data in the SWIR range, discusses spectral characteristics of spectra from a burn model, and shows that unsupervised classification methods such as *k*-means could be applied for burn classification.

In the field of forensics, HSI has been used to investigate crime scenes, to identify stains of biological fluids and blood on soft materials such as clothing, and for characterizing skin bruises.[73,74,106–108] Bruises are caused by ruptured blood vessels and leakage of blood into the tissue. Immediately after the trauma, there will be a wheal and flare reaction, which is characterized by swelling and fluid accumulation.[109] It has been shown that, in the SWIR spectral range, it is possible to differentiate between blood and other fluids accumulating in the injured area[74] (Figs. 5.6 and 5.7 show the spectra from bruised skin in subjects with skin types II and IV). These spectra clearly show that the influence from pigmentation is limited in the SWIR spectral range. Paintballs induce circular bruises with an undamaged central spot, due to central compression of the skin directly underneath the paintball, and the shear force at the edge of the paintball causes vessel rupture. When imaging a fresh paintball bruise in the SWIR, the central area will appear swollen, while a rapidly developing hemorrhage can be seen at the edges.

5.3 Hyperspectral Instrumentation and Setup

HSI was initially developed for remote sensing and satellite imaging. In 2005, Sellar and Boreman classified imaging spectrometers for HSI[34] and defined a taxonomy for imaging spectrometers. According to Boreman's definitions, and highlighted in an overview provided by Wu and Sun,[13] hyperspectral instrumentation and data collection is divided into four types. The first type is

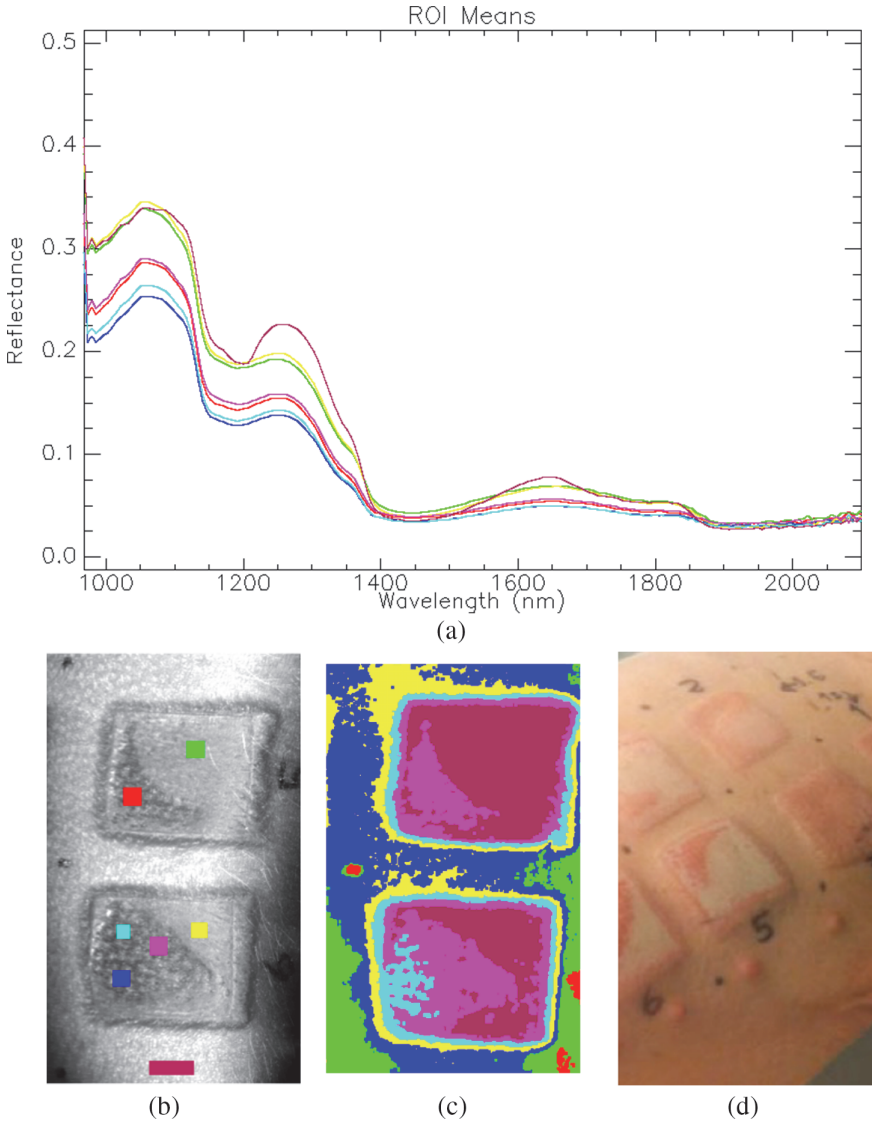

Figure 5.5 Hyperspectral burn severity assessment. (a) Spectra from the locations shown in part (b). (b) Locations for spectra shown in part (a). (c) Burn severity classification using unsupervised k-means classification. (d) Photograph of the induced burn injuries. [Data and part (d) from Paluchowski et al.[53] Parts (a), (b), and (c) reprinted from Randeberg.[75]]

whisk broom scanning, which operates by first collecting all of the wavelengths from one pixel, then repeating the process and scanning—pixel by pixel—to collect spatial data. The second type is push broom scanning, where all wavelengths and spatial information in one row of spatial pixels are collected simultaneously, then the scene is scanned line by line. The third type is wavelength scanning, where the entire scene is collected one wavelength at

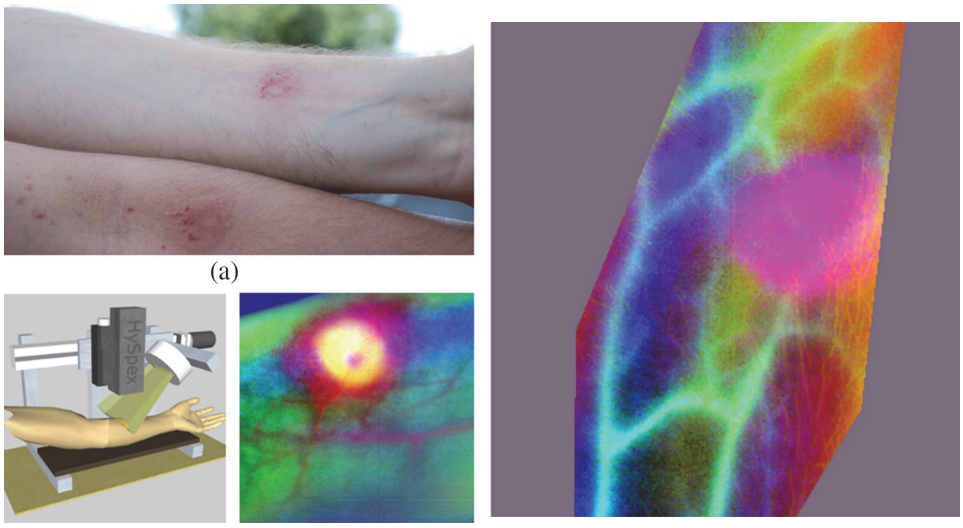

Figure 5.6 Hyperspectral visualization of paintball bruises in the SWIR spectral range. (a) Photo of paintball injuries shortly after infliction. (b) HSI setup consisting of a camera and a light source attached to a translation stage. (c) Synthetic RBG image constructed from the first three bands of a minimum noise fraction (MNF) transform of a SWIR hyperspectral image of a fresh paintball injury. The pink color indicates hemorrhaging, and the yellow indicates accumulating tissue fluid. (d) Synthetic RBG image constructed from the first three bands of a MNF transform of a SWIR hyperspectral image of a paint ball bruise a few days after infliction. The pink area is the location of the bruise. (Adapted from Randeberg and Hernandez-Palacios.[74])

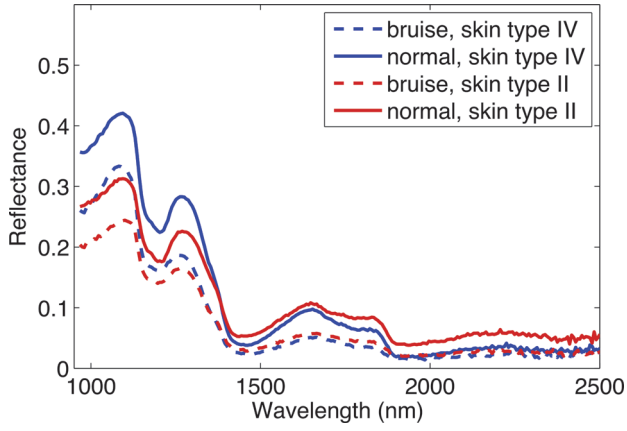

Figure 5.7 Reflectance spectra derived from bruised skin in volunteers having Fitzpatrick skin types II and IV. (Reprinted from Randeberg and Hernandez-Palacios.[96])

the time, then the wavelengths are scanned. In this case, the wavelength scanning might take place on either the sensor or illumination side of the system. The fourth type is snapshot imaging, where both spectral and spatial coordinates are collected at the same time. These scanning modes are illustrated in Fig. 5.8.

Most hyperspectral scanning systems use one of these basic scanning configurations, even if they are combined with other imaging modalities or optical systems. All of these configurations have strengths and weaknesses. In summary, pixel scanning and push broom systems are vulnerable to motion artifacts between pixels/lines. Effects such as breathing and movement might affect the spatial image, but not the spectral data. Wavelength-scanning devices might suffer from movement artifacts between the wavelength bands but not between the spatial pixels. The snapshot imagers should not in

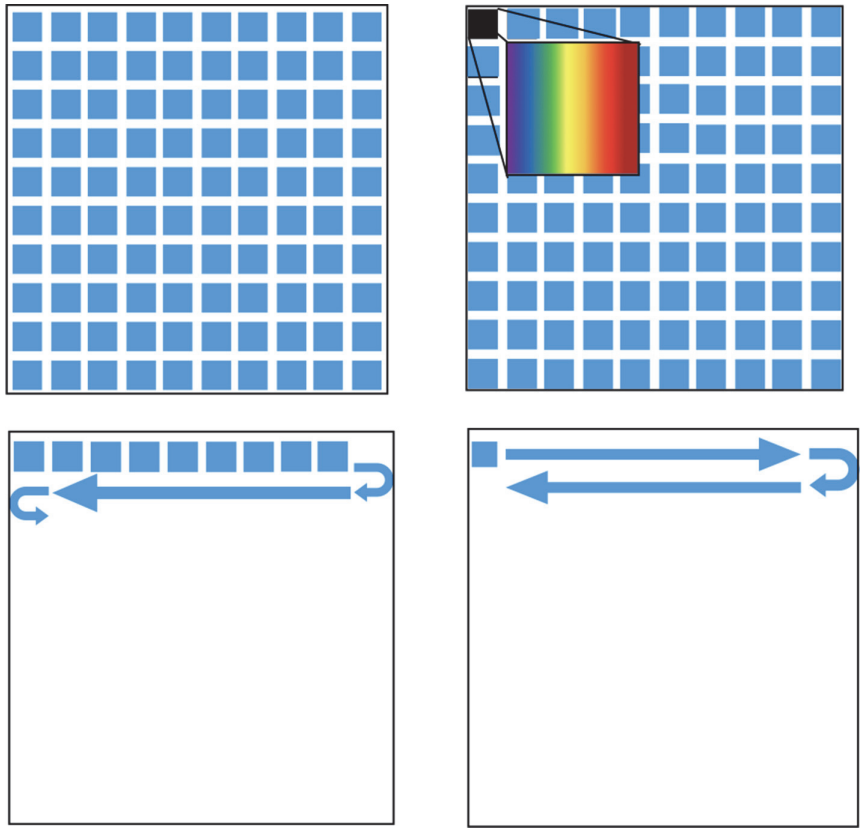

Figure 5.8 Scanning modes. Top left: snapshot, collecting all the spatial and spectral information from all pixels at the same time; top right: wavelength scanning, collecting all spatial information at the same time, but scanning the wavelengths; lower left: push broom, collecting all information for one spatial row of pixels, and scanning row by row; lower right: whisk broom, spatially scanning pixel by pixel, collecting all spectral data for each pixel.

principle be prone to movement artifacts but may have other issues, such as spectral resolution and spectral-band density.

Several detector technologies are relevant for the SWIR spectral range. Technical systems used in HSI and medicine are described in Refs. 1, 13, and 25. According to Adão,[25] indium arsenide (InAs), gallium arsenide (GaAs), indium gallium arsenide (InGaAs), and mercury cadmium telluride (MCT or HgCdTe) are the dominant semiconductor materials used as detectors in the SWIR range (refer to Fig. 5.9 for further details). In contrast, silicon-based detectors are used for VIS/NIR imaging with wavelengths up to approximately 1100 nm. However, silicon-based detectors cannot be used with infrared light due to the bandgap of the material (i.e., silicon-based detectors are not sensitive to photons with wavelengths longer than 1100 nm). The processes for manufacturing SWIR detectors are more demanding and complex, and the sensors are typically more expensive and of lower specifications than VIS/NIR counterparts.

Experimentally, there are two different, basic configurations for hyperspectral data acquisition: light transmission and light reflection (Fig. 5.10). For materials that are semitransparent to one or more wavelengths in the SWIR range, and with a sufficiently thin thickness, a measurement to collect the light transmitted through the sample can be implemented. In this modality, a light source emitting photons of adequate wavelengths is placed on one side of the sample, and the camera is placed on the other side of the sample. The camera

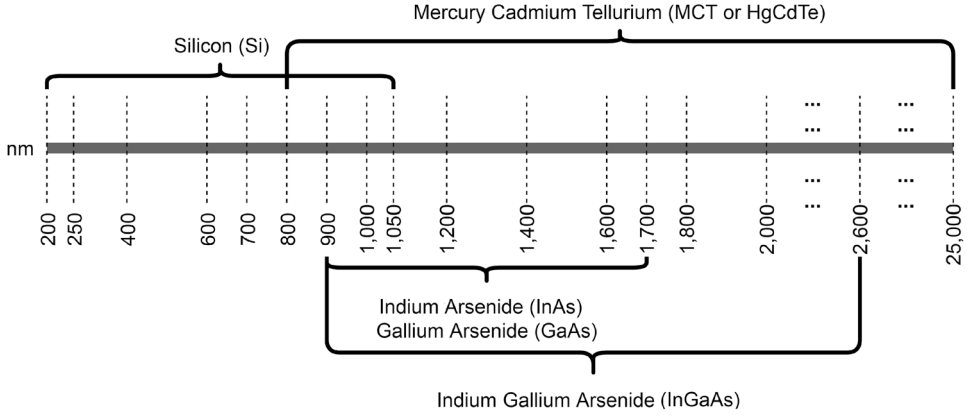

Figure 5.9 Materials involved in hyperspectral sensor fabrication: Si is used for acquiring UV, VIS, and shortwave NIR regions; InAs and GaAs have a spectral response between 900 and 1700 nm; InGaAs and GaAs have a spectral response between 900 and 1700 nm (InGaAs extends the previous range to 2600 nm, and MCT in InGaAs extends the previous range to 2600 nm); and MCT is characterized by a large spectral range and high quantum efficiency that enables reaching the MWIR region (approximately 2500–5000 nm) and NIR region (approximately 800–2500 nm). [Reprinted from Adão et al.[25] Published under a creative commons 4.0 license.]

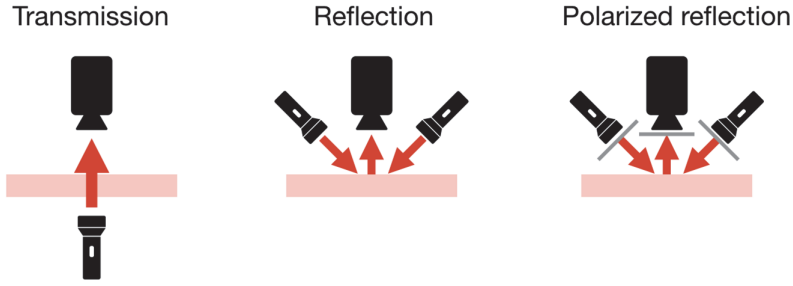

Figure 5.10 Measurement configurations.

then captures the photons that emerge on the other side of the sample after interacting with the sample, revealing information on the material.

However, in reflection mode, the light source and the camera are placed on the same side of the sample. As the light impinges on the surface of the sample, the signal is reflected either specularly or diffusely, depending on the texture and materials present on the surface. To a lesser extent, there can be a contribution to the reflected signal from the light that penetrates the sample and is backscattered to the surface. Usually, the reflected signal collected by the hyperspectral sensor is a combination of all of these interactions.

In cases where the specular reflection dominates, the reflected signals can complicate interpretation of the data; the system can be set up in a cross-polarized configuration, where the angle of the incoming light plus polarizers on the light sources and on the camera can be used to remove or reduce the surface glare.[14] This configuration, or a variant involving other states of polarization, can also be implemented to investigate materials and structures with a reflectivity response affected by the polarization of light. It should be noted that the polarizers used for this type of assessment are also wavelength dependent: in the same way that detectors in the VIS/NIR region are not suitable for the SWIR range, polarizers for the VIS/NIR such as those used in photography and standard microscopy may not be suitable for polarizing light in the SWIR.

An additional imaging modality known as transflection is a combination of transmission and reflection. This mode, also termed transillumination or interactance, utilizes a light source to illuminate a section of a sample, while the hyperspectral sensor is focused on a different region—away from that in which the light impinges. The signal, which is measured by the camera as the light that penetrates the sample, diffuses in the direction of the section being imaged—and is then backscattered to the surface of the sample.

Point scanning and push broom HSI systems, which have a dispersive element to gather the spectral information, need strong illumination to return proper signal-to-noise ratios (SNRs), as the light is distributed into wavelength bins on the detector. The choice of light source must match the

camera system, optics, measurement geometry, and sample. Choice of illumination is thus one of the most important decisions to ensure a successful measurement. Biological samples can become damaged by high irradiances from the light source. The scanning geometry, scanning time, and nature of the sample determine the highest advisable irradiance. In the SWIR, halogen lamps are still frequently used, as well as lasers and light-emitting diodes (LEDs). Halogen lamps tend to have a high infrared contribution, creating extended problems with drying and heating of the samples. If possible, the lamps should be fitted with infrared filters to remove heat. During all work with live animals or humans, International Commission on Nonionizing Radiation Protection (ICNIRP) guidelines for exposure should be kept. These guidelines are valid for lasers, LEDs, and other continuous sources.

One of the main shortcomings of biomedical HSI is the limited optical penetration depth in tissue. Even though the penetration depth is substantially larger in the SWIR spectral region, it is challenging to acquire good data with a sufficient SNR. This is especially challenging with thick tissues and large samples. Due to the complexity of the optical system, focusing of a hyperspectral system can be challenging. Focusing depends on the measurement configuration and the depth of focus of the optics.

In microscopic and endoscopic systems, focusing is handled by the endoscope or microscope optics, while in a wide-field system, closeup lenses of variable- or fixed-focus distance and different aperture sizes can be used. A hyperspectral camera using adjustable optics (lenses, typically C-mount-based, with variable aperture size and working distance, and attached directly to the entrance aperture of the camera) allows for manual or automatic adjustment of focus. Although flexible and relatively inexpensive, adjustable closeup lenses introduce optical aberrations to the system that may impact the data quality.[110]

The optical elements in the camera have transmissive and reflective properties that are wavelength dependent. The light traveling through the optical system may follow different paths or have different magnifications or depths of focus, depending on the wavelengths of the incoming signal. While the most common aberrations in a traditional optical system (e.g., astigmatism, chromatic aberration, coma, distortion, field curvature, and spherical aberration in scientific photography) can be reduced or eliminated using high-end lenses, HSI systems are affected differently, as the necessary corrections must act over the entire spectral range in very narrow bandwidths. The most substantial distortions affecting imaging spectrometers are the keystone and smile effects. Keystone is nonuniform, wavelength-dependent magnification that effectively introduces a spatial misregistration of the light signal from a scene pixel on the detector. On the other hand, the smile effect is a nonuniform shift of the center of gravity of the spectral bands along the field of view (FOV) of the camera; it is responsible for spectral misregistration of the image from the instantaneous

FOV of the camera. These effects are present in some form in most hyperspectral cameras, and a distinction between high-end and low-end systems can be made based on the degree to which they affect the quality of the data. Adding focusing lenses to a given camera will most likely increase the deviations from both types of aberrations, especially if the lenses are not designed for the system. The selection of lenses for an application should consider the potential repercussions on the distortions, and the user should keep in mind that spatial sharpness does not ensure spectral fidelity.

Depending on the depth of field of the lens, the differences in height across the sample may pose a challenge for achieving a sharp image. In other words, when imaging samples with a varying topography at the scale of the depth of field or larger, some regions may appear sharp while others could be out of focus. A method to overcome such challenges could be to collect a stack of images with different focuses and combine the images in postprocessing. It is also important to notice that spatial sharpness does not necessarily ensure spectral sharpness. The optical elements in the camera have transmissive and reflective properties that are wavelength dependent. Light traveling through the optical system may follow different paths and have different magnifications and depths of focus—depending on the wavelengths of the incoming signal.

Alternatively, the hyperspectral system may be fitted with a fixed-focus lens designed to limit the effects of optical aberrations at the expense of a limited focal range. When designed carefully, custom closeup lenses may be used to maintain the spectral fidelity of the imaging system while working at different focal distances, magnifications, and *f*-number values. The depth of field will also limit the usability as previously described, and layered imaging can be implemented by moving the camera away from and closer to the sample to change the focusing plane of the camera. For more-complex configurations, the process of adjusting the working distance of the system to scan the different levels of the focal plane can be automated by combining the hyperspectral scanning with a working distance scanning following, e.g., a trajectory along a three-dimensional (3D) surface model of the sample.[111]

5.4 Hyperspectral Data Collection

Selection of the imaging mode, camera specifications, and illumination source depend on the application of interest. With a setup in place, the data acquisition can be a straightforward process, but several considerations must be taken into account. Push broom imagers offer high spatial resolution and high spectral band density, and can be used in a variety of configurations. For information on multispectral snapshot cameras or Fourier-transform hyperspectral imagers, refer to, e.g., Cao et al.[112] or Su and Sun.[113]

The most critical parameter for data acquisition is the integration time (IT), which is equivalent to the exposure time in traditional photography. The IT defines the time during which the sensor collects the light signal coming from the sample at a given position of the scanning process. After this period, the sensor transfers the collected data to an internal readout module, resets the pixels in the detector, moves to the next position of the scanning, i.e., the next line of the image, and repeats the light acquisition process until the entire sample has been imaged. In high-sensitivity cameras, typical values for the IT are on the order of milliseconds. All detectors used in hyperspectral sensors have intrinsic noise parameters (e.g., electronic, thermal, and dark current) that are typically low for high-end instruments but are not negligible, especially in the SWIR range. The SNR is a measure of the expected quality of an image and, given that there is noise intrinsic to the sensor, one way to increase the SNR is by increasing the signal intensity, which is a function of the IT. An alternative way to increase the SNR without increasing the IT is by using a more powerful light source to illuminate the samples. However, because biological tissue is susceptible to damage by light of high intensity, this approach is not always feasible. Another option to increase the SNR consists of scanning the same line in an image multiple times and averaging the signal. This reduces the photon noise but increases the duration of the acquisition. Finally, the noise in an image should also be quantified for each scan such that the data can be corrected during postprocessing. An image of the dark signal, i.e., the image recorded by the camera when no light is coming through the sample, should be part of every image acquisition. Some cameras require the user to manually cover the entrance aperture of the camera, while others do this automatically using an internal shutter.

Another operational parameter of relevance for image acquisition is the scanning speed. The speed at which the sample or the camera moves must be set such that the geometry of the sample is accurately imaged by the sensor. If the scanning speed is too high, i.e., the IT is longer than the time it takes the scanning system to cover a distance equivalent to the lateral size of one scene pixel, then the data in each pixel will be undersampled and the image will look compressed. In contrast, if the speed is too slow, the image will be stretched, and the data size will be higher than at slower speeds. A simple test to validate the speed used for a given IT consists of imaging objects of known geometrical proportions and verifying that the measured size corresponds to the imaged size in two dimensions.

Setting the illumination in the optimal position is also an important task of image acquisition. The light sources used may be of different types (as mentioned in Section 5.3) and may use lenses, waveguides, or optical fibers to direct the light onto the samples. It is important to arrange the sources in such a way that the FOV of the camera is uniformly illuminated. The variations on the topography of the samples may create shadowed areas in the

scene that will have a lower SNR; this should be avoided whenever possible by optimizing the orientation and distribution of the illumination sources.

Finally, whenever the data is to be compared to other existing data or used to define spectral libraries to be used as references, it should be possible to transform the acquired data from the native format of the camera to absolute reflectance, absorbance, or transmittance values. For this transformation, it is necessary to process the data using a reference pattern of known reflectivity, absorbance, or transmissivity as a function of wavelength.

5.5 Hyperspectral Data Analysis

A hyperspectral image may contain hundreds of wavelengths, where each wavelength can be considered as a separate image and each pixel as a separate spectrum. Thus, the images contain more information than it is possible for the human brain to process. Hyperspectral data can be analyzed using different methods; the high spectral resolution requires methods from spectroscopy and chemometrics, while the spatial resolution requires methods from computer vision. Figure 5.11 shows a typical workflow in hyperspectral data analysis. Basically, the processing chain starts with the raw datacube, which is then preprocessed into the desired form before image analysis is applied to identify a structure or signature, or for quantification or classification.

Calibration of hyperspectral images was already covered in a previous section, and dimensionality reduction, noise removal, and other preprocessing tools are described in Section 5.7. From a biomedical optics perspective, light transport simulations have been an important tool; this aspect is covered in Section 5.6. Machine learning methods and chemometric tools are covered in Section 5.7, and deep learning methods in Section 5.8.

5.6 Data Analysis and Simulations

Knowledge of the optical properties of tissue can reveal important information, including whether the tissue is benign or malignant. This information can be obtained using spectroscopic data and mathematical

Figure 5.11 Typical workflow for hyperspectral processing. If deep learning is applied, some of this work might be reduced.

modeling. There are numerous studies on modeling the optical properties of tissues—which can depend on the available data—using the VIS/NIR spectral range, while there are fewer studies on modeling in the SWIR region. This lack of data limits use of optical modeling (e.g., Monte Carlo modeling or diffusion theory), although the availability and number of datasets have improved.[56,84–86] Spectroscopic studies and modeling have facilitated more-extensive modeling using SWIR-HSI.[18,93,114,115] Zhang et al.[115] explore the penetration depth of light in biological tissues from 900–1650 nm using Monte Carlo simulations and a HSI system, with the Michelson spatial contrast as a metric of light penetration. Wang et al.[93] present a method to estimate full-field burn depth based on HSI in the wavelength range of 950–1650 nm. They employed a method based on regression models to obtain the estimate, but also present Monte Carlo data on estimated light penetration in burns. Milanic et al.[116] used Monte Carlo simulations to estimate the transmittance and reflectance from human hands in order to investigate if HSI could be used to diagnose arthritis. They found that there is an optical window of high transmission in the vicinity of 1100 nm, where the penetration is at its maximum and there is minimal influence of water and melanin absorption (Fig. 5.12). Wrobel et al.[18] simulated a noise-free hyperspectral Fourier-transform infrared imaging dataset of a pancreatic tissue core based on experimental data to test the performance of a selection of processing algorithms in a controlled manner. The simulated dataset allowed for controlled addition of noise with a given characteristic.

Modeling can also be utilized in interpretation and unmixing of data from turbid and thick samples. In a high-scattering tissue, the reflectance measured at the surface will contain information from structures and tissues located at various depths into the tissue. This is also the case in the SWIR spectral range. The surface signals will be modulated due to the scattering properties of the sample; this modulation might complicate use of standard machine learning and chemometric tools to detect and quantify the composition of a biological sample. Application of machine learning methods will require use of scattering correction models in turbid media; as machine learning increases in popularity, physics-based methods might find a new role in hybrid approaches where such methods are used both for scattering corrections and to explain the physical and biological context behind the results.

5.7 Chemometric Tools and Methods from Spectroscopy

Chemometric data processing can be utilized in spectroscopy (i.e., Raman and infrared)[117] to reveal information on variables that may be intercorrelated within the spectra.[118,119] Using chemometrics, it is possible to disclose the relationships among the variables and efficiently identify patterns in the data. Chemometric methods are based on multivariate analysis, which exploits the

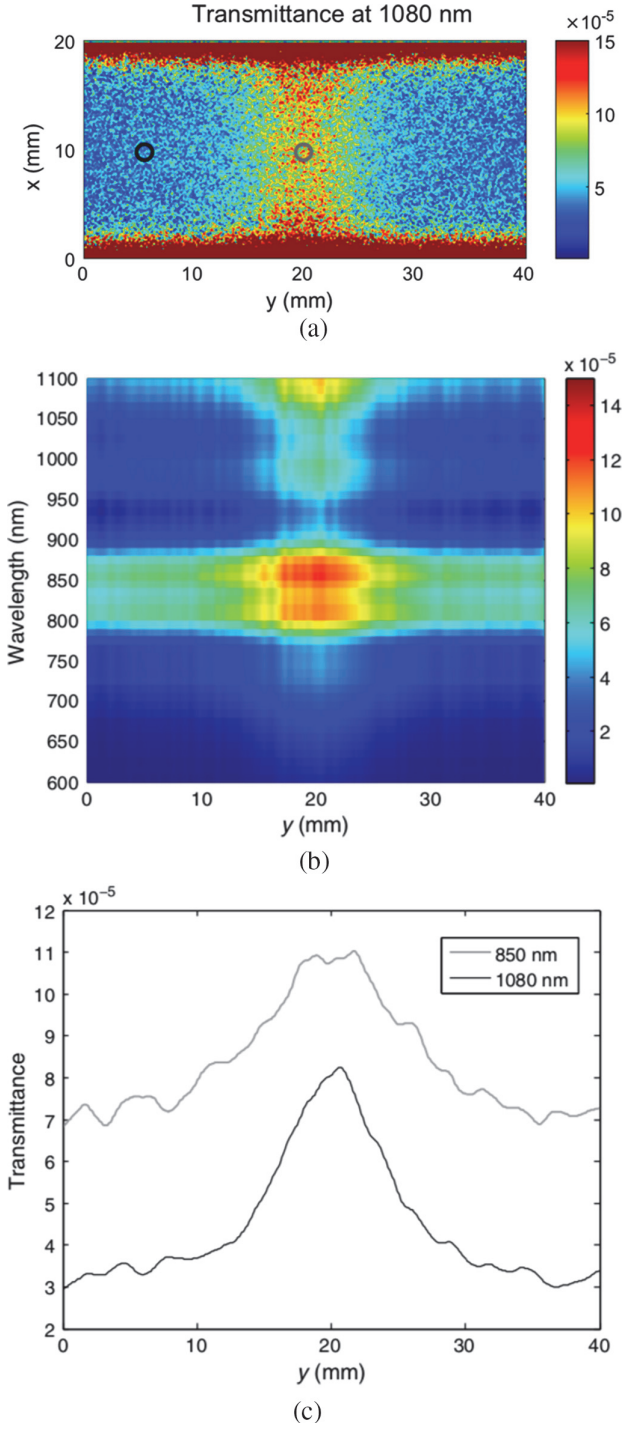

Figure 5.12 Monte Carlo simulation of light transmission through a human finger. (a) Transmission through the joint at 1080 nm. (b) Transmission along the finger. (c) Transmission through a joint at 850 and 1080 nm. [Figures reprinted from Milanic et al.[116]]

intercorrelation of the data to build a more straightforward and condensed representation of the information contained therein.[24] Chemometrics is especially developed to extract the chemical information contained in complex data and is applied in many fields such as medicine,[120] pharmaceutics,[121,122] food science,[24] and forensics.[123]

Among the various spectroscopic techniques, SWIR spectroscopy in the wavelength range of 900–2500 nm has substantially benefited from chemometric methods because the signals—such as vibrational features, overtones, and combination bands—usually overlap and are difficult to interpret.[124,125] Preprocessing of the data is a required and essential step to remove spurious information and unwanted variations.[126] Conventionally, chemometric methods and tools are applied to matrices of data with dimensions ($N \times l$), where N represents the samples or objects (observations) organized in rows, and l represents the variables (wavelengths) organized in columns.[127]

Over the last 10 years, chemometrics has moved with ease from a two-dimensional (2D) dataset ($N \times l$) to a 3D dataset ($N \times M \times l$) after introduction of HSI technologies. Corresponding applications of multivariate image analysis have effectively contributed to interpretation of data by obtaining images that show the distribution of chemical (or physical) properties of the system under study.[128] The majority of chemometric algorithms can efficiently work on 3D data by reorganizing the 3D datacube into a 2D matrix. This procedure is called unfolding and leads to a matrix reorganized with pixels in the rows and wavelengths in the columns. The critical point is to unfold the data following a particular order such that it is possible to reinstate each pixel in the correct position and display the results of the processed data [i.e., by principal component analysis or cluster analysis] in the original image space[128] (Fig. 5.13).

Chemometric methods are basically machine learning methods and can be divided into unsupervised and supervised methods. Unsupervised methods are aimed at exploring the data and disclosing patterns naturally present in the dataset without a defined ground truth to aid analysis. Conversely, supervised methods require *a priori* knowledge of the collected data and help to make predictions of new variables or to collocate them in defined classes.[129] The following sections discuss the most common unsupervised and supervised methods applied to analyze spectral data, focusing on SWIR spectroscopy and HSI applications.

5.7.1 Spectral preprocessing

Preprocessing is a fundamental step in spectroscopy that is taken before applying any chemometric method. First, preprocessing can remove undesired variation in the data,[129] leaving mainly the chemical information. Second, it can enhance weak, but important, signals by redistributing the variables' weight.[124,130] Spectral preprocessing is usually applied before unsupervised or

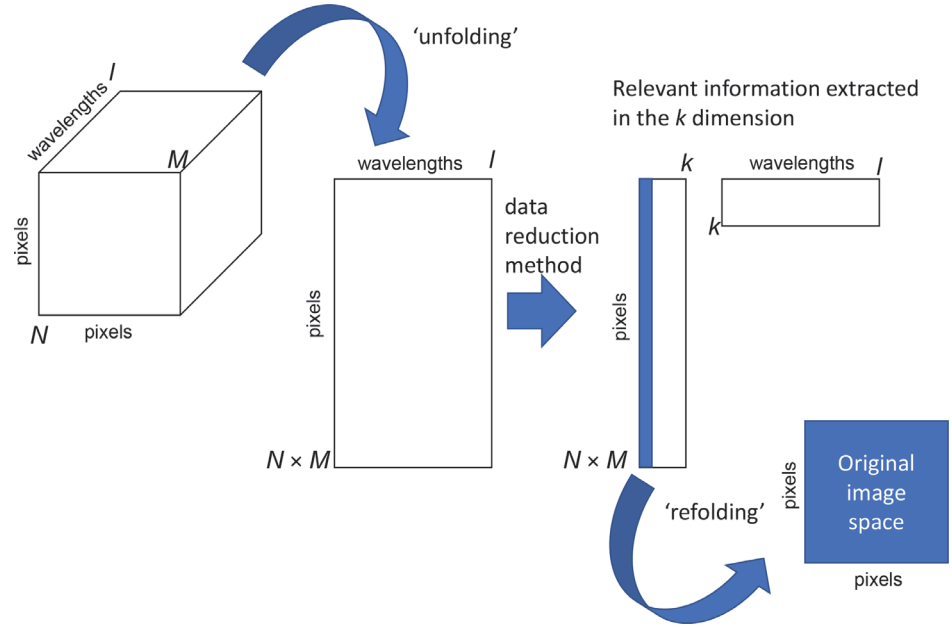

Figure 5.13 Procedure followed by chemometrics to operate on hyperspectral data. The datacube is unfolded in a 2D data matrix.

supervised methods, as it has proved to be capable of improving the performance of chemometric models in terms of interpretability and prediction, and/or classification properties.[131] However, preprocessing methods should be used carefully as they might cause misleading results.[132]

All of the commonly known preprocessing methods from spectroscopy can be easily extended to HSI, as the procedure is applied along the spectral dimension.[133]

Noise removal is often the first preprocessing step to be applied to spectroscopic data. Noise can be removed in various ways, using different filtering methods. The most common of these uses smoothing filters, where a moving window—simple or combined with a fitting polynomial (i.e., the Savitzky–Golay[134] method)—is moved along the signal to be filtered.[126]

Scattering effects are common and may be the cause of the nonlinear behavior of the data, especially in the NIR/SWIR region.[135] The effect of scattering is usually evident when spectra show vertical baseline shifts or different intensity effects in the bands, i.e., general band amplification or decrease.[132,135] If the desire is to extract chemical information, the scattering effects should be removed, as they cause high variability. Multiplicative scatter corrections and standard normal variate (SNV) can be successfully applied for this purpose. Scattering can also cause—together with instrumental factors and spectroscopy effects (i.e., fluorescence)—distortion of the spectral baseline. As baseline distortion can lead to the loss of linearity in the

spectra, it is advisable to correct the spectral shape before applying chemometric methods. Usually, baseline correction is performed with detrending methods, where a trend in the spectra is modeled as a polynormal and successively subtracted from the data.[126] Scattering removal can also be carried out using physics-informed methods and photon transfer modeling.[136]

Mean centering and scaling are common preprocessing operations that occur before multivariate data analysis. Mean centering brings the data to the center of the coordinate system, simplifying the interpretation of the multivariate model.[130] Data scaling is useful as it allows one to readjust the importance of the variables, especially in the presence of heterogeneous parameters collected from different techniques. Scaling can also be used to enhance the contribution of specific variables that are important for construction of the model.[126]

Use of derivative filters is a common but complicated filtering method that can be applied as preprocessing. The main advantage of performing a derivative is that it allows for—in a simple operation—the removal of vertical shifts or both shifts and drifts, depending on whether the first or second derivative is applied.[132] Moreover, it is possible to highlight spectral features[133] or resolve and visualize overlapping bands in complex spectral profiles, bringing emphasis to the smallest differences in the data.[126] A problem in applying derivatives is noise, which is usually amplified together with the signal. In this case, the derivation should be undertaken using the Savitzky–Golay method, which implies a preliminary smoothing step that improves the output of the derivative signal.[134]

5.7.2 Unsupervised methods

5.7.2.1 Principal component analysis

The most well-known unsupervised technique for exploratory data analysis is principal component analysis (PCA). PCA reduces the data into a set of a few uncorrelated orthogonal variables [called latent variables (LVs) or principal components (PCs)] that contain the relevant information of the dataset. The new variables capture the data's variance such that the first variable describes the largest variance, followed by the other variables in decreasing order of importance. PCA decomposes the matrix $\mathbf{X}$ following a bilinear model, which can be mathematically represented by[137]

$$\mathbf{X} = \mathbf{T}\mathbf{P}^{\mathrm{T}} + \mathbf{E}, \tag{5.1}$$

where $\mathbf{X}$ ($N \times l$) is the preprocessed (mean-centered, scaled, and so on) data matrix, T ($N \times A$) is the matrix of scores, $\mathbf{P}$ ($l \times A$) is the matrix of loadings (superscript T denotes transpose), and $\mathbf{E}$ ($N \times l$) is the matrix of residuals. The value of A is the optimum number of PCs that best captures the relevant information of the data. Each column in the score matrix $\mathbf{T_A}$ (and each

column in the loading matrix $\mathbf{P_A}$) captures the variability of a single PC that is uncorrelated with the variability of the other columns or PCs. The scores are the coordinates of the new variables in the PC space, while the loadings represent the importance of each original variable for each PC.[137]

A useful way to interpret information in the scores and loadings is to create 2D or 3D scatter plots, where each PC is plotted against one other. Score plots usually highlight similarities/dissimilarity of samples, revealing the presence of clusters or specific trends in the data. Loading plots, inspected jointly with score plots, using the same pair of PCs, allow for identifying the most relevant variables responsible for forming the clusters in the scores, with the variables pointing to each corresponding cluster.[138]

Score and loading plots have been widely exploited in spectroscopy to find naturally occurring structures associated with similarities and dissimilarities in the spectral signals. Additionally, by plotting each loading vector as a function of wavelength, a spectral-like visualization of the loadings can be obtained—known as loading profiles. Loading profiles can be used to interpret the data because a high positive or negative band intensity denotes meaningful wavelengths. Positive loadings are associated with with positive scores, while negative loadings are associated with negative scores.[139] However, bands with negative values (due to the loadings' orthogonality property) may sometimes increase the difficulty in interpreting the loading profiles.[138]

PCA applied to hyperspectral images also gives score and loading plots and loading profiles. Since, in this case, we are dealing with thousands or even millions of objects (spectral pixels), it is useful to inspect scatter plots where the pixel density is represented. In this way, identification of agglomerates and their distribution in the PC space is immediate [Fig. 5.14 (b)]. In some cases, it is interesting to know whether a particular cluster in the score plot corresponds to a specific segment in the original image. This information can be obtained by segmentation, where it is possible to manually select the pixels clusters in the score plots and have them highlighted with a color in the original image by using dedicated tools such as brushing [Figs. 5.14(b) and (d)].[140] This tool is useful as it allows one to group and visualize pixel spectra that may be clustered by similar spectral shapes and possibly chemical properties.[128,138] The reverse possibility, namely, selecting a segment of the image and visualizing its correspondence in the score plot, is also possible. A useful improvement of this tool allows for extracting the averaged spectrum from the selected cluster of pixels, rendering it even more straightforward to identify spectral profiles that are characteristic of specific areas.[141,142]

Another possibility allowed by PCA on hyperspectral images is to represent the information retained by each score vector in the form of an image. Each score vector has the number of rows equal to the total number of pixels of the image (Fig. 5.13) and can be conveniently refolded into the

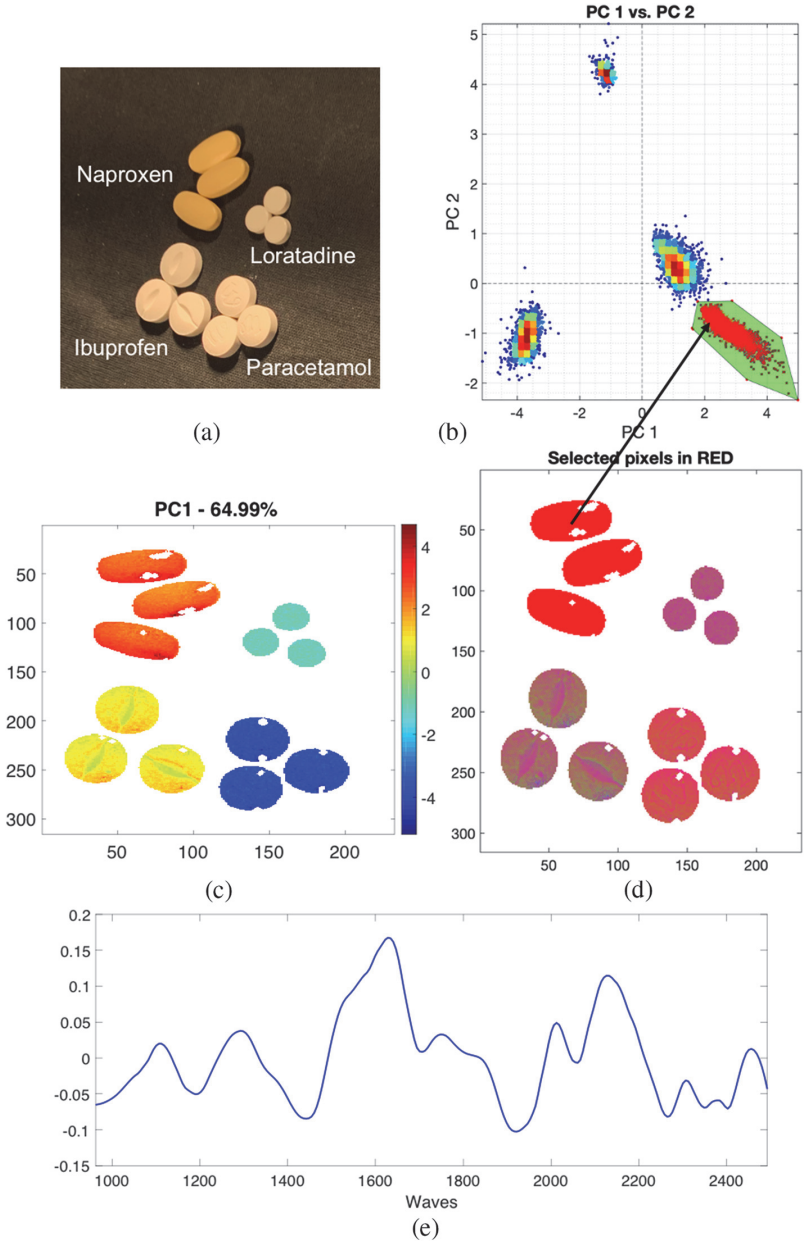

Figure 5.14 Common tools for inspecting data after applying PCA[140] (data were preprocessed using Savitzky–Golay smoothing and SNV). (a) RGB photo of the four pharmaceuticals that were scanned (naproxen, paracetamol, ibuprofen, and loratadine). (b) Score plot PC1 versus PC2, where a particular cluster (red) has been selected. (c) PC1 score image (64.99%). (d) Highlighted pixels (red) corresponding to the selected pixels in the score plot. (e) Loading profile for PC1. [Data were processed using the software Hypertools (https://www.hypertools.org/). Hyperspectral data were collected using a Hyspex SWIR 320-me camera (Hyspex, Oslo, Norway).]

original image space ($N \times M$), preserving the original pixel position. This representation is particularly convenient for enhancing the similarities and dissimilarities between pixels (similarities that are possibly linked to similar chemical properties) for each PC. A grayscale or false-color code can represent each pixel's score values and provide an intuitive visualization of pixel similarities in the score images.

A critical problem related to PCA is selecting the optimal number A of PCs. There are several methods, but generally the selection is made by inspecting the eigenvalues or residual variance plot against the number of PCs, known as a scree plot. The scree plot usually shows a line with a large negative slope for the PCs that contain relevant information, and then becomes flat where the PCs do not contain any more useful information. Researchers base their choice of how many PCs to retain according to the PC value that indicates a marked change in the slope of the curve, keeping all of the PCs before that selected point.

Besides the possibilities offered by the aforementioned tools to clarify PCA results, the dimensionality reduction capability offered by this technique represents an enormous advantage. Data compression performed by PCA can facilitate more-informative data processing methods applied only to the relevant information. It should be mentioned that several other filtering techniques such as minimum noise fraction (MNF) or independent component analysis (ICA) also can be used for dimensionality reduction and feature extraction.

5.7.2.2 Cluster analysis

Unsupervised cluster analysis is designed to group objects according to the similarity of their spectral properties based on the information retained in the data. Objects that belong to a particular group share similar features, whereas they differ from other groups/objects.[143] In spectroscopy, cluster analysis is applied to group spectra with similar spectral profiles, which share common chemical properties or information. Cluster analysis has been widely used to search for interesting patterns or as a starting point for further, more-complex data processing. The family of clustering methods is wide and sometimes confusing, but a well-established distinction is made, which determines whether clustering methods are hierarchically, partitionally, or density-based.[143,144] Hierarchical clustering progressively groups objects according to a given definition of similarity or dissimilarity. The common representation of this agglomerative-based (bottom-up) approach is a cluster tree or dendrogram, which shows at which level the objects and/or groups are merged. In the opposite way (top-down), the data structure is considered as one cluster and is progressively split into smaller clusters.[143]

Alternatively, partitional clustering divides the dataset into a defined number of clusters, where each object is allocated to only a single cluster.

Partitional clustering uses an iterative algorithm to optimize certain parameters (i.e., the centroids of the clusters) and to obtain the best membership of objects to a cluster, as well as the estimated number of clusters. Finally, density-based methods cluster objects by looking preferentially at the high-density regions that characterize them. A density threshold and the volume to be inspected around each object are the two critical parameters that must be set from the beginning. Afterward, the method proceeds automatically to find the optimal number of clusters.[144]

Since similarity is the key point of clustering, an important step is defining a set of metrics that can measure the distance between the objects and between clusters of objects. The distance measures between objects (x) can be summarized with the general Minkowski formula:

$$d_{ij} = \left[\sum_{k=1}^{N} (x_{ik} - x_{jk})^{\alpha}\right]^{1/\alpha}, \tag{5.2}$$

where d_{ij} is the distance between object i and j. When $\alpha = 1$, the measure is the Manhattan distance, and when $\alpha = 2$, the measure is Euclidean.

Several methods have been proposed to measure distances between clusters, considering the previous formula; examples are single linkage, complete linkage, and group average (unweighted pair group methods). Single-linkage groups clusters according to the smallest distance between their closest objects. Conversely, complete linkage uses the smallest distance between the most distant objects. Group average uses the small distance between the centroids. Hierarchical clustering makes extensive use of distances between clusters, as it proceeds by measuring and grouping clusters until only one cluster is reached. Hierarchical clustering has been applied in medicine to cluster data characterized by different properties.[143]

One of the most commonly used clustering methods is k-means clustering, which is unsupervised; i.e., the number of k clusters should be given by the user before starting. The algorithm will find the centers of k clusters and iteratively improve their positions until a stable situation is achieved and the centroids no longer move. The k-means technique is a hard-clustering method, meaning that each object is unequivocally assigned to only one cluster. There is also a fuzzy version of k-means, which assigns the objects a fraction of cluster membership represented by a value between 0 and 1. An object can thus belong simultaneously to different clusters, and the clusters can overlap. This is a suitable method to use in applied sciences as there may be some overlap between clusters. A problem in using k-means is determining the number of clusters, which is not usually known in advance.[143] Several methods have been developed to find the most suitable number of clusters. One interesting method was proposed by Hastie et al.,[145] where different values of k are tested, and the correspondent within-sum of squares for each k

is plotted against the number of clusters. An abrupt slope change in the sum of squares curve may be used as a marker to identify the optimum number of clusters.

Cluster analysis has been applied to hyperspectral data in fields such as pharmaceutics[146] and remote sensing.[147] Since cluster analysis can easily process many pixels but decreases in performance with a large number of variables, useful results can be achieved by applying cluster analysis to multivariate images where a small number of latent variables (images) are created by using PCA, MNF, or similar transforms.[138,144] This has been applied to SWIR hyperspectral in medicine, where MNF was performed before applying *k*-means clustering of data.[75] As the image representation of the results is important in cluster analysis, there are several ways to bidimensionally display the results. In *k*-means, the output vector—which contains the cluster integer for every pixel—can be easily refolded into a matrix **C** with the same dimensions as the original image. In this case, every pixel's membership can be viewed with a color in the original image space.

5.7.2.3 Methods for purest spectra extraction

A common problem in spectroscopy and HSI is identification and isolation of spectral signatures that are unique to a single compound. This is difficult to achieve since spectra collected from different samples are often a mixture of different compounds. Separating (i.e., resolving) the mixed spectral signals—in pure spectral profiles and relative concentrations—is a challenging field of research, and several methods have been proposed.[148,149] Among other methods, self-modeling mixture analysis approaches are unsupervised techniques aimed at finding—without any previous knowledge of the mixture—the pure/purest variables and their concentration profiles in a series of measured spectra. By definition, a pure variable accounts for one component and has an intensity proportional to its concentration. For example, consider a series of SWIR spectra organized in a matrix **D** of dimensions $N \times l$, where N represents the spectra, and l represents the wavelengths. Finding the pure variables with self-modeling mixture analysis approaches involves operating along the columns of **D** and extracting only the variables 1446, 1710, and 1945 nm, each one accounting specifically for only one component. To each variable an associated concentration profile is given. Finding the pure variables is interesting in all branches of spectroscopy, as these variables can eventually be used to calculate the spectra of pure compounds in the mixture.[150] Nevertheless, having a single pure variable instead of a spectrum might also be a bit limiting, especially for spectroscopies that use the simultaneous presence of several characteristic bands to correctly characterize a compound.

An interesting solution to this problem is offered by the same aforementioned self-modeling methods by applying the selection method

along the rows instead of the columns of the matrix **D**. In this case, we obtain a set of spectra extracted directly from the original data matrix, with each spectrum containing the highest contribution of one component. These spectra rarely correspond to the spectra of pure compounds since bands of other compounds might also be present, but they represent the best set of purest spectra for the compounds present in the mixture. Extraction of the purest spectra may offer a valid contribution in spectral interpretation as the user analyzes a spectrum that contains the most substantial spectral bands of a specific compound.

The main idea that unites the self-modeling techniques is that the purest spectra (or pure variables) are the most diverse, especially in their corresponding mixtures. The most common techniques for extracting the purest spectra (or pure variables) are key set factor analysis (KSFA), orthogonal projection approach (OPA), and simple-to-use interactive self-modeling mixture analysis (SIMPLISMA).[150] All of these techniques share some common steps, namely, (1) an initial estimation of the first purest spectrum (or variable) and (2) use of the determinant to find the most-diverse spectra from the first and successive spectra (or variables) selected. When the purest spectra (variables) are individuated, their row (column) position is determined and easily recovered from the original matrix. Besides these general common steps, the algorithms were developed quite differently by following their own criteria.

Before initiating the algorithms, the number of purest spectra to extract should be determined. This value is usually obtained by calculating the original matrix's rank (typically by PCA or SVD), considering that each meaningful component (large variance) will be equivalent to a purest spectrum of the diverse substances contained in the mixture. Here we briefly describe the three aforementioned self-modeling techniques, with an emphasis on the problem of extracting the purest spectral profiles.

KSFA is a technique developed by Malinowski[151] that searches the minimum spectra set that best characterizes the dataset. In this method, the purest spectra are identified based on the concept that the purest are the most diverse and thus the most orthogonal. To achieve this, PCA is first applied to the data matrix to obtain score and loading matrices. Second, the score vectors (or loading if we are interested in pure variables) are normalized to the unit length (as we are mostly interested in the vectors' directions and not their magnitudes). To find the first purest spectrum, the smallest absolute value along the first score vector (column) is considered. Its row position corresponds to the position of the first purest spectrum in the original data matrix. To find the second purest spectrum, a matrix **C** of dimensions $N \times 2$, which contains the first and second score vectors, is considered. Now two score values $[v_{a1}\ v_{a2}]$ from the row of the first purest spectrum are extracted and compared with all pairs of score values taken along the rows r of the

C matrix. For each combination, a 2 × 2 matrix **M** is constructed, and the absolute value of the determinant of this matrix is computed as

$$A = |\det(M)| = \left|\det\begin{bmatrix} v_{a1} & v_{a2} \\ v_{j1} & v_{j2} \end{bmatrix}\right|, \tag{5.3}$$

where $[v_{a1}\, v_{a2}]$ are the two score values of the matrix **C** corresponding to the first purest spectrum, while $[v_{j1}\, v_{j2}]$, with j = number of rows of the matrix, are all other pairs of score values of the reduced-score **C** matrix. The determinant measures the degree of orthogonality between the set of values and can find the most-diverse values (spectra) from the highest determinant. The pairs of score values that maximize the determinant will define a new row position, which corresponds to the second purest spectrum in the original matrix. The algorithm proceeds by increasing the number of score vectors (for component 3, **C** is of dimensions $N \times 3$) and inserting new rows in the determinant calculation; then, again, the new spectrum is selected when the maximum determinant is reached. The algorithm stops when the k rows corresponding to the number of k purest spectra selected by the user are reached. Schostack and Malinoski[152] propose an iterative development of the previous KSFA algorithm to solve the problem of possible lack of orthogonality in the set of spectra. In the iterative version of KSFA (IKSFA), the determinant is recalculated by iteratively replacing the key rows with all of the key rows not tested in the previous algorithm. If the determinant is improved during one of the substitutions, the new key set replaces the old one. The main steps in KSFA are presented in Table 5.1.

IKSFA was applied mainly in the field of chromatography to extract pure chromatographic profiles.[153] However, in the field of vibrational spectroscopy/HSI, very few applications exist; these are mainly reported by Catelli et al.,[154] where IKSFA has been used to extract the set of purest spectra from SWIR-HSI data collected from the surface of a highly corroded bronze

Table 5.1 Main steps in KSFA.

Step number	Step Description
1	Analysis of matrix **X** by PCA
2	Normalization of score vectors (columns)
3	Individuation of the first purest spectrum from the absolute value of the lowest score
4	Calculation of the absolute value of the determinant of pairs of score vectors taken along the rows, using two columns of the score matrix
5	Identification of the pairs of vectors that gave the highest determinant
6	Individuation of the second purest spectrum
7	Repeat of steps 4–6 in increments of 1 for the number of score vectors along the rows and columns
8	Halt of the analysis when the k number of the purest spectra is obtained

sculpture. The extracted purest spectra were interpreted, then the spatial distribution of the corresponding compounds were displaced using a spectral angle mapper tool.[154]

The OPA developed by Cuesta Sánchez et al.[155] calculates the purest spectra based on the highest dissimilarity principle among the purest spectra. A reference spectrum is initially computed by taking the mean spectrum, normalized to a length equal to 1. The first purest spectrum is obtained by finding the spectrum that has the highest dissimilarity with the mean. To this aim, the mean spectrum is combined iteratively with all other spectra in the matrix inside a matrix **F** of dimensions $2 \times l$, where the first row contains the mean spectrum and the second row contains, sequentially, all other spectra. The dissimilarity between the two spectra is obtained by calculating the determinant of the correlation matrix of **F**, $\det(\mathbf{F}\mathbf{F}^{\mathrm{T}})$. The candidate spectrum that gives the highest determinant is the most dissimilar and is selected as the first purest spectrum. In a second step, the first purest spectrum is used as a reference spectrum, and the previous procedure is repeated to calculate the most dissimilar spectrum, which gives the second purest spectrum. All of the purest spectra found are included at each new step in matrix **F**, and the aforementioned procedure is repeated until the k values are reached.

In SIMPLISMA,[150,156] the position of the purest spectra in the original matrix is obtained by calculating the purity value p_i. The purity value is computed using two common statistical tools—the mean μ and standard deviation σ:

$$p_i = \frac{\sigma_i}{\mu_i + \alpha}, \tag{5.4}$$

where i is the number of the spectra. Estimation of noise is included in the calculation of the purity value (offset value = α) to compensate for the fact that high purity might be given to a variable that consists of only noise. Initially, the row with the highest p_i values corresponds to the row with the first purest spectrum. Several subsequent steps are taken to calculate the second purest spectrum, but these steps are basically designed to remove everything related to the first purest spectrum. In practice, this is achieved by scaling the original data matrix and then by computing the correlation matrix $\mathbf{D} = \mathbf{Z}\mathbf{Z}^{\mathrm{T}}$ of a pair of spectra (**Z** is a $2 \times l$ matrix), where one spectrum is the first purest spectrum, and the other spectrum is selected iteratively inside the scaled matrix. The highest value of the determinant of **D** identifies the second purest spectrum. Similar to OPA, the procedure is again iterated to find the set of k purest spectra. The determinant values are also called weights w, and the values will be close to zero for those rows related to the purest spectra but will be highest for rows with no relation to the purest.[150] Thus, the purity parameter for each purest spectrum at each step can be rewritten as

$$p_i = w_i \frac{\sigma_i}{\mu_i + \alpha}. \tag{5.5}$$

OPA and SIMPLISMA were largely applied in the fields of spectroscopy[157,158] and HSI,[159] often combined with advanced multivariate curve resolution methods.[160–162] Recently, several alternatives to SIMPLISMA were also applied to hyperspectral data.[163]

5.7.3 Supervised method: regression

Regression analysis comprises a series of techniques that aim to find a relationship between a variable of y, called the response or dependent variable, and other variables x, called predictors or independent variables. Finding the best relationship between y and x variables means that we must find the best values of unknown coefficients b, associated with the predictors, which satisfy the following equation:

$$y = b_1 x_1 + b_2 x_2 + \ldots + b_M x_M + e. \tag{5.6}$$

The b coefficients are called regression coefficients and can be estimated from recorded data using the least-squares method in a common regression problem.[164] The variable y is the response, and e is the residual. In the spectroscopy field, Eq. (5.6) may represent the possibility of relating the concentration of an analyte y and its corresponding spectrum (with x variables). If multiple measurements of concentrations and spectra are produced, Eq. (5.6) can be rewritten in a condensed matrix form:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}, \tag{5.7}$$

where $\mathbf{y}$ is an $(N \times 1)$ vector, $\mathbf{X}$ is an $(N \times l)$ matrix with N spectra and l wavelengths, $\mathbf{b}$ is an $(l \times 1)$ vector of coefficients, and $\mathbf{e}$ is an $(N \times 1)$ vector. Our goal is to determine the best values of $\mathbf{b}$ to construct a relationship (model) that can best predict future $\mathbf{y}$ values, starting from freshly acquired variables $\mathbf{X}$ (spectra). This is also known as a calibration problem. The reason for the wide range of regression methods is that the variables $\mathbf{X}$ are easy or straightforward to acquire (such as spectroscopic measurements), while the variables $\mathbf{y}$ are difficult and complicated to obtain (i.e., the concentration of analytes). The aforementioned method, called multiple linear regression, shows some limitations if the number of samples is lower than the number of variables and the variables are highly correlated, which are common features of spectroscopy.[139]

Multivariate regression methods were developed to solve the regression problem by using multivariate analysis. Multivariate methods are generally applied when we want to find a relationship between $\mathbf{y}$ and a high number of variables in $\mathbf{X}$. In this way, instead of using the full $\mathbf{X}$ matrix, we construct a

regression model exploiting a small number of uncorrelated latent variables that capture the best variability of the data. This was quite an improvement, especially in spectroscopy, as it is possible to solve the common problem of the usually high number of variables compared to the samples.[165]

The most common multivariate regression methods are principal component regression (PCR) and partial least-squares regression (PLSR). PCR finds the best LVs for **X**, and their directions are used to construct the regression model for **y**. However, the direction of the PCs in **X** might not represent the highly relevant information for prediction of **y**. For this reason, PLSR, or simply PLS, was introduced and elaborated on to find the LVs in **X** that best represent the variability in **y**. Thus, the multivariate modeling was transformed into a more-efficient predictive ability of the technique. The PLS method is usually divided into PLS1, where the **y** matrix has only one column (one response per measurement), and PLS2, where the **Y** matrix has more than one column (multiple responses per measurement). The PLS2 model generally is useful when there is a correlation between the **Y** variables. A key point in PCR and PLS is determining the correct number of LVs or PLS components. This is usually obtained by inspecting the prediction error [root mean-square error prediction (RMSEP)] calculated on predicted y_p and experimentally sampled y_e values, starting from a new matrix $\mathbf{X_e}$ of the following values:

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n} (y_e - y_p)^2}{n}}. \tag{5.8}$$

The prediction error plot as a function of the number of PLS components follows a U shape in which the minimum corresponds to the optimal number of PLS components. If we exceed the number of components, we have an overfitted model or a model strongly dependent on the specific properties of the data used to construct the model. If the number of components is too low, we have an underfitted model that loses some of the data's important information.[166]

A fundamental step in regression is the model's validation, which is necessary to estimate the model's future predictive ability. We refer to $\mathbf{X}_{\text{cal}}$ and $\mathbf{y}_{\text{cal}}$ as the matrices used to obtain regression parameters (calibration model), while $\mathbf{X}_{\text{val}}$ and $\mathbf{y}_{\text{val}}$ are the matrices of validation. The best method to validate a model is using an independent test set, meaning an extensive group of new $\mathbf{X}_{\text{val}}$ and $\mathbf{y}_{\text{val}}$. The data set used to build the calibration model is known as a training set, while the data used for the validation is known as a test set. However, there is often not enough data for the validation test, as the acquisition of y variables is rather complex. Therefore, the test set should be constructed internally from the calibration set by splitting it into training and test sets. In such a case, we need to resample from the original dataset to create a test set of objects. Among the most common resampling techniques,

cross-validation (CV) temporarily removes part of the data to simulate an independent test set.[166]

With either a new validation test set or an independent validation test set, we can evaluate the results of the calibration model using two equations: the root mean-square error (RMSE) equation and the R^2 coefficient, both of which are terms for the calibration, cross-validation, and predictive ability. Low values of the RMSE with an R^2 coefficient close to 1 indicate good performances of the model in terms of predictive ability:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}} \tag{5.9}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y}_i)^2}, \tag{5.10}$$

where y and $\bar{y}$ (mean) are experimental, $\hat{y}$ is the predicted value for sample i, and n is the total number of samples.

5.8 Machine Learning and Artificial Intelligence

It is no doubt that machine learning [a form of artificial intelligence (AI)] and especially deep learning can greatly enhance HSI. By using appropriate auto-encoders, the algorithms can work on data without dimensionality reduction, noise removal, or separate feature extraction. However, there are disadvantages. In 2018, a workshop was held to develop a roadmap for foundational research on AI in medical imaging.[167] Researchers highlighted several state-of-the-art applications of AI in medicine, which are as follows:

1. New image reconstruction methods that efficiently produce images suitable for human interpretation from source data.
2. Automated image labeling and annotation methods—including information extraction from the imaging report, electronic phenotyping, and prospective structured image reporting.
3. New machine-learning methods for clinical imaging data—such as tailored, pretrained model architectures, and federated machine learning methods.
4. Machine-learning methods that can explain the advice they provide to human users (i.e., explainable AI).

5. Validated methods for image de-identification and data sharing to facilitate wide availability of clinical imaging data sets.

The increase in the accuracy and performance of AI systems—especially deep neural networks—comes with more-complex models, rendering them harder to interpret.[168,169] Recent studies have shown that an explanation of the model is important and allows the user to trust algorithmic advice and decisions.[170–172] Currently, the main focus is on adapting existing algorithms and developing new algorithms to achieve segmentation, classification, detection, and quantification. However, typically, AI requires large and almost unrealistic amounts of data to ensure a smooth performance. The lack of sufficient data hinders full utilization of the strength of AI for data analysis. This requires use of techniques such as data augmentation and transfer learning to strengthen the dataset for training. Hyperspectral data are complex and deep, but not necessarily big in terms of numbers. Another challenge in the biomedical field is the lack of an established ground truth for training algorithms. Quite often there are no other comparable techniques or established benchmarks, implying that the ground truth depends on human observation or selection—which is inherently subjective and uncertain. An error of 20% per human observation will propagate through the system and result in incorrectly trained algorithms. Unsupervised clustering can be a valuable technique in cases where it is challenging to use or interpret results from physics-based models, or where the ground truth is lacking or not well defined.

Despite all of these drawbacks, tasks such as classification, detection, and segmentation have seen tremendous progress.[52] Khan et al.[52] discuss different types of neural networks and deep learning methods, and give examples on how these approaches have been applied to medical HSI. Halicek et al.[21] present an overview of analyses applied in cancer studies. However, this overview is a mixture of deep learning and more-traditional machine learning.

Trajanovski et al.[167] claim "the first study using deep learning semantic segmentation for tumor detection in HSI data using channel selection, and accounting for more spatial tissue context, and global comparison between the prediction map, and the annotation per sample. Accurate deep learning algorithms have a huge potential to be a promising alternative to digital pathology or a doctors' supportive tool in real-time surgeries." Their results show good segmentation of cancerous and normal tissue. In general, it seems that cancer research is ahead of the rest of the biomedical community with respect to adapting AI into research; de Koning et al.[48] and Baltussen et al.[81] also applied such methods.[1] Readers who seek a well-written, concise introduction to the possibilities for applying deep learning to hyperspectral data should read Audebert et al.[174]

Acknowledgments

Thanks to I. A. Randeberg for assistance with Fig. 5.3.

References

1. G. Lu and B. Fei, "Medical hyperspectral imaging: a review," *J. Biomed. Opt.* **19**(1), 010901 (2014) [doi: 10.1117/1.JBO.19.1.010901].
2. M. Manley, "Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis of biological materials," *Chem. Soc. Rev.* **43**(24), 8200–8214 (2014).
3. J. A. Gutiérrez-Gutiérrez, A. Pardo, E. Real, J. M. López-Higuera, and O. M. Conde, "Custom scanning hyperspectral imaging system for biomedical applications: modeling, benchmarking and specifications," *Sensors* **19**(7), 1692 (2019).
4. R. Fakhrullin, L. Nigamatzyanova, and G. Fakhrullina, "Dark-field/hyperspectral microscopy for detecting nanoscale particles in environmental nanotoxicology research," *Sci. Total Environ.* **772**, 145478 (2021).
5. S. Ortega, M. Halicek, H. Fabelo, G. M. Callico, and B. Fei, "Hyperspectral and multispectral imaging in digital and computational pathology: a systematic review," *Biomed. Opt. Express* **11**(6), 3195–3233 (2020).
6. S. Lemmens, J. Van Eijgen, K. Van Keer, J. Jacob, S. Moylett, L. De Groef, T. Vancraenendonck, P. De Boever, and I. Stalmans, "Hyperspectral imaging and the retina: worth the wave?," *Transl. Vis. Sci. Technol.* **9**(9), 9 (2020).
7. G. Saiko, P. Lombardi, Y. Au, D. Queen, D. Armstrong, and K. Harding, "Hyperspectral imaging in wound care: a systematic review," *Int. Wound J.* **17**(6), 1840–1856 (2020).
8. J. Shapey, Y. Xie, E. Nabavi, R. Bradford, S. R. Saeed, S. Ourselin, and T. Vercauteren, "Intraoperative multispectral and hyperspectral label-free imaging: a systematic review of in vivo clinical studies," *J. Biophotonics* **12**(9), e201800455 (2019).
9. L. Gao and R. T. Smith, "Optical hyperspectral imaging in microscopy and spectroscopy—a review of data acquisition," *J. Biophotonics* **8**(6), 441–456 (2015).
10. A. U. Rehman and S. A. Qureshi, "A review of the medical hyperspectral imaging systems and unmixing algorithms in biological tissues," *Photodiagnosis Photodyn. Ther.* **33**, 102165 (2020).
11. N. Mehta, S. Shaik, R. Devireddy, and M. R. Gartia, "Single-cell analysis using hyperspectral imaging modalities," *J. Biomech. Eng.* **140**(2), 020802 (2018).
12. S. Ortega, H. Fabelo, D. K. Iakovidis, A. Koulaouzidis, and G. M. Callico, "Use of hyperspectral/multispectral imaging in

gastroenterology. Shedding some different light into the dark," *J. Clin. Med.* **8**(1), 36 (2019).
13. D. Wu and D.-W. Sun, "Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: a review—part I: fundamentals," *Innovative Food Sci. Emerging Technol.* **19**, 1–14 (2013).
14. E. L. Larsen, L. L. Randeberg, A. Aksnes, L. O. Svaasand, E. Olstad, and O. A. Haugen, "Hyperspectral imaging of atherosclerotic plaques in vitro," *J. Biomed. Opt.* **16**(2), 026011 (2011) [doi: 10.1117/1.3540657].
15. N. Mehta, S. P. Sahu, S. Shaik, R. Devireddy, and M. R. Gartia, "Dark-field hyperspectral imaging for label free detection of nano-bio-materials," *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **13**(1), e1661 (2021).
16. Y. Zhang, S. Zhu, J. Lin, and P. Jin, "High-quality panchromatic image acquisition method for snapshot hyperspectral imaging Fourier transform spectrometer," *Opt. Express* **27**(20), 28915–28928 (2019).
17. V. Notarstefano, S. Sabbatini, C. Conti, M. Pisani, P. Astolfi, C. Pro, C. Rubini, L. Vaccari, and E. Giorgini, "Investigation of human pancreatic cancer tissues by Fourier transform infrared hyperspectral imaging," *J. Biophotonics* **13**(4), e201960071 (2020).
18. T. P. Wrobel, P. Koziol, M. K. Raczkowska, D. Liberda, C. Paluszkiewicz, and W. M. Kwiatek, "Noise-free simulation of an FT-IR imaging hyperspectral dataset of pancreatic biopsy core bound by experiment," *Sci. Data* **6**(1), 239 (2019).
19. A. A. Gowen, Y. Feng, E. Gaston, and V. Valdramidis, "Recent applications of hyperspectral imaging in microbiology," *Talanta* **137**, 43–54 (2015).
20. J. Yoon, J. Joseph, D. J. Waterhouse, A. S. Luthman, G. S. D. Gordon, M. Di Pietro, W. Januszewicz, R. C. Fitzgerald, and S. E. Bohndiek, "A clinically translatable hyperspectral endoscopy (HySE) system for imaging the gastrointestinal tract," *Nat. Commun.* **10**(1), 1902 (2019).
21. M. Halicek, H. Fabelo, S. Ortega, G. M. Callico, and B. Fei, "In-vivo and ex-vivo tissue analysis through hyperspectral imaging techniques: revealing the invisible features of cancer," *Cancers* **11**(6), 756 (2019).
22. K. Okubo, "NIR hyperspectral imaging," in *Transparency in Biology*, K. Soga, M. Umezawa, and K. Okubo, Eds., Springer, Singapore, pp. 203–222 (2021).
23. L. A. Paluchowski, E. Misimi, L. Grimsmo, and L. L. Randeberg, "Towards automated sorting of Atlantic cod (Gadus morhua) roe, milt, and liver—spectral characterization and classification using visible and near-infrared hyperspectral imaging," *Food Control* **62**, 337–345 (2016).
24. P. Oliveri, C. Malegori, and M. Casale, "Chemometrics: multivariate analysis of chemical data," in *Chemical Analysis of Food: Techniques and*

Applications, 2nd edition, Academic Press, Cambridge, Massachusetts, 33–76 (2020).
25. T. Adão, J. Hruška, L. Pádua, J. Bessa, E. Peres, R. Morais, and J. J. Sousa, "Hyperspectral imaging: a review on UAV-based sensors, data processing and applications for agriculture and forestry," *Remote Sens.* **9**(11), 1110 (2017).
26. I. Pope, F. Masia, K. Ewan, A. Jimenez-Pascual, T. C. Dale, F. A. Siebzehnrubl, P. Borri, and W. Langbein, "Identifying subpopulations in multicellular systems by quantitative chemical imaging using label-free hyperspectral CARS microscopy," *Analyst* **146**(7), 2277–2291 (2021).
27. K. Bae, L. Xin, W. Zheng, C. Tang, B.-T. Ang, and Z. Huang, "Mapping the intratumoral heterogeneity in glioblastomas with hyperspectral stimulated Raman scattering microscopy," *Anal. Chem.* **93**(4), 2377–2384 (2021).
28. E. Lipiec, J. Kaderli, J. Kobierski, R. Riek, K. Skirlińska – Nosek, K. Sofińska, M. Szymoński, and R. Zenobi, "Nanoscale hyperspectral imaging of amyloid secondary structures in liquid," *Angew. Chem. Int. Ed.* **60**(9), 4545–4550 (2021).
29. C. Di Napoli, I. Pope, F. Masia, P. Watson, W. Langbein, and P. Borri, "Hyperspectral and differential CARS microscopy for quantitative chemical imaging in human adipocytes," *Biomed. Opt. Express* **5**(5), 1378–1390 (2014).
30. N. Billecke, G. Rago, M. Bosma, G. Eijkel, A. Gemmink, P. Leproux, G. Huss, P. Schrauwen, M. K. Hesselink, M. Bonn, and S. H. Parekh, "Chemical imaging of lipid droplets in muscle tissues using hyperspectral coherent Raman microscopy," *Histochem. Cell Biol.* **141**(3), 263–273 (2014).
31. C. H. Camp, J. S. Bender, and Y. J. Lee, "Real-time and high-throughput Raman signal extraction and processing in CARS hyperspectral imaging," *Opt. Express* **28**(14), 20422–20437 (2020).
32. I. Zorin, R. Su, B. Heise, B. Lendl, and M. Brandstetter, "Correlative infrared optical coherence tomography and hyperspectral chemical imaging," *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **37**(9), B19–B26 (2020).
33. M. B. Applegate, S. S. Spink, and D. Roblyer, "Dual-DMD hyperspectral spatial frequency domain imaging (SFDI) using dispersed broadband illumination with a demonstration of blood stain spectral monitoring," *Biomed. Opt. Express* **12**(1), 676–688 (2021).
34. R. G. Sellar and G. D. Boreman, "Classification of imaging spectrometers for remote sensing applications," *Opt. Eng.* **44**(1), 013602 (2005) [doi: 10.1117/1.1813441].
35. R. Koprowski and P. Olczyk, "Segmentation in dermatological hyperspectral images: dedicated methods," *Biomed. Eng. Online* **15**(1), 97 (2016).

36. G. Florimbi, H. Fabelo, E. Torti, R. Lazcano, D. Madroñal, S. Ortega, R. Salvador, F. Leporati, G. Danese, A. Báez-Quevedo, and G. M. Callicó, "Accelerating the K-nearest neighbors filtering algorithm to optimize the real-time classification of human brain tumor in hyperspectral images," *Sensors* **18**(7), 2314 (2018).
37. Y. Khouj, J. Dawson, J. Coad, and L. Vona-Davis, "Hyperspectral imaging and K-means classification for histologic evaluation of ductal carcinoma in situ," *Front. Oncol.* **8**, 17 (2018).
38. W. Markgraf, P. Feistel, C. Thiele, and H. Malberg, "Algorithms for mapping kidney tissue oxygenation during normothermic machine perfusion using hyperspectral imaging," *Biomed. Eng.* **63**(5), 557–566 (2018).
39. J. Wang and Q. Li, "Quantitative analysis of liver tumors at different stages using microscopic hyperspectral imaging technology," *J. Biomed. Opt.* **23**(10), 106002 (2018) [doi: 10.1117/1.JBO.23.10.106002].
40. E. Kho, L. L. de Boer, K. K. Van de Vijver, F. van Duijnhoven, M. J. T. V. Peeters, H. J. Sterenborg, and T. J. Ruers, "Hyperspectral imaging for resection margin assessment during cancer surgery," *Clin. Cancer Res.* **25**(12), 3572–3580 (2019).
41. J. B. Lamy, B. Sekar, G. Guezennec, J. Bouaud, and B. Séroussi, "Explainable artificial intelligence for breast cancer: a visual case-based reasoning approach," *Artif. Intell. Med.* **94**, 42–53 (2019).
42. L. Li, W. Li, Y. Qu, C. Zhao, R. Tao, and Q. Du, "Prior-based tensor approximation for anomaly detection in hyperspectral imagery," *Proc. IEEE Trans. Neural Netw. Learn. Syst.* **2020**, 1–14 (2020).
43. D. Sato, T. Takamatsu, M. Umezawa, Y. Kitagawa, K. Maeda, N. Hosokawa, K. Okubo, M. Kamimura, T. Kadota, T. Akimoto, and T. Kinoshita, "Distinction of surgically resected gastrointestinal stromal tumor by near-infrared hyperspectral imaging," *Sci. Rep.* **10**(1), 21852 (2020).
44. F. Manni, F. van der Sommen, H. Fabelo, S. Zinger, C. Shan, E. Edström, A. Elmi-Terander, S. Ortega, and G. M. Callicó, "Hyperspectral imaging for glioblastoma surgery: improving tumor identification using a deep spectral-spatial approach," *Sensors* **20**(23), 6955 (2020).
45. F. Manni, R. Fonolla, F. V. der Sommen, S. Zinger, C. Shan, E. Kho, S. B. de Koning, T. Ruers, and P. H. N. de With, "Hyperspectral imaging for colon cancer classification in surgical specimens: towards optical biopsy during image-guided surgery," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* **2020**, 1169–1173 (2020).
46. S. Ortega, M. Halicek, H. Fabelo, R. Guerra, C. Lopez, M. Lejeune, F. Godtliebsen, G. M. Callico, and B. Fei, "Hyperspectral imaging and deep learning for the detection of breast cancer cells in digitized

histological images," *Proc. SPIE* **11320**, 113200V (2020) [doi: 10.1117/12.2548609].
47. M. Halicek, J. D. Dormer, J. V. Little, A. Y Chen, and B. Fei, "Tumor detection of the thyroid and salivary glands using hyperspectral imaging and deep learning," *Biomed. Opt. Express* **11**(3), 1383–1400 (2020).
48. S. G. B. de Koning, P. Weijtmans, M. B. Karakullukcu, C. Shan, E. J. M. Baltussen, L. A. Smit, R. L. P. van Veen, B. H. W. Hendriks, H. J. C. M. Sterenborg, and T. J. M. Ruers, "Toward assessment of resection margins using hyperspectral diffuse reflection imaging (400–1,700 nm) during tongue cancer surgery," *Lasers Surg. Med.* **52**(6), 496–502 (2020).
49. R. Wang, Y. He, C. Yao, S. Wang, Y. Xue, Z. Zhang, J. Wang, and X. Liu, "Classification and segmentation of hyperspectral data of hepatocellular carcinoma samples using 1-D convolutional neural network," *Cytom. Part A* **97**(1), 31–38 (2020).
50. B. Jansen-Winkeln, M. Barberio, C. Chalopin, K. Schierle, M. Diana, H. Köhler, I. Gockel, and M. Maktabi, "Feedforward artificial neural network-based colorectal cancer detection using hyperspectral imaging: a step towards automatic optical biopsy," *Cancers* **13**(5), 967 (2021).
51. D. Saha and A. Manickavasagan, "Machine learning techniques for analysis of hyperspectral images to determine quality of food products: a review," *Curr. Res. Food Sci.* **4**, 28–44 (2021).
52. U. Khan, S. Paheding, C. Elkin, and V. Devabhaktuni, "Trends in deep learning for medical hyperspectral image analysis," *IEEE Access* **9**, 79534–79548 (2021).
53. L. A. Paluchowski, H. B. Nordgaard, A. Bjorgan, H. Hov, S. M. Berget, and L. L. Randeberg, "Can spectral-spatial image segmentation be used to discriminate experimental burn wounds?," *J. Biomed. Opt.* **21**(10), 101413 (2016) [doi: 10.1117/1.JBO.21.10.101413].
54. A. Bjorgan, B. Pukstad, and L. Randeberg, "Hyperspectral characterization of re-epithelialization in an in vitro wound model," *J. Biophotonics* **13**(10), e202000108 (2020).
55. A. Bjorgan and L. Randeberg, "Exploiting scale-invariance: a top layer targeted inverse model for hyperspectral images of wounds," *Biomed. Opt. Express* **11**(9), 5070–5091 (2020).
56. R. H. Wilson, K. P. Nadeau, F. B. Jaworski, B. J. Tromberg, and A. J. Durkin, "Review of short-wave infrared spectroscopy and imaging methods for biological tissue characterization," *J. Biomed. Opt.* **20**(3), 030901 (2015) [doi: 10.1117/1.JBO.20.3.030901].
57. D. Boas, C. Pitris, and N. Ramanujam, Eds., *Handbook of Biomedical Optics*, CRC Press, Boca Raton, Florida (2011).
58. I. J. Bigio and S. Fantini, *Quantitative Biomedical Optics: Theory, Methods and Applications*, Cambridge University Press, Cambridge, United Kingdom (2016).

59. R. Goodacre, E. M. Timmins, R. Burton, N. Kaderbhai, A. M. Woodward, D. B. Kell, and P. J. Rooney, "Rapid identification of urinary tract infection bacteria using hyperspectral whole-organism fingerprinting and artificial neural networks," *Microbiology* **144**(5), 1157–1170 (1998).
60. J. A. Timlin, A. Carden, M. D. Morris, J. F. Bonadio, C. E. Hoffler, K. Kozloff, and S. A. Goldstein, "Spatial distribution of phosphate species in mature and newly generated mammalian bone by hyperspectral Raman imaging," *J. Biomed. Opt.* **4**(1), 28–34 (1999) [doi: 10.1117/1.429918].
61. M. F. Parker, G. C. Mooradian, J. P. Karins, D. M. O'Connor, B. A. Speer, P. D. Owensby, and A. Velasco, "Hyperspectral diagnostic imaging of the cervix: report on a new investigational device," *J. Low. Genit. Tract Dis.* **4**(3), 119–124 (2000).
62. D. G. Ferris, R. A. Lawhead, E. D. Dickman, N. Holtzapple, J. A. Miller, S. Grogan, S. Bambot, A. Agrawal, and M. L. Faupel, "Multimodal hyperspectral imaging for the noninvasive diagnosis of cervical neoplasia," *J. Low. Genit. Tract Dis.* **5**(2), 65–72 (2001).
63. K. J. Zuzak, M. D. Schaeberle, M. T. Gladwin, R. O. Cannon, III, and I. W. Levin, "Noninvasive determination of spatially resolved and time-resolved tissue perfusion in humans during nitric oxide inhibition and inhalation by use of a visible-reflectance hyperspectral imaging technique," *Circulation* **104**(24), 29052910 (2001).
64. K. J. Zuzak, M. D. Schaeberle, E. N. Lewis, and I. W. Levin, "Visible reflectance hyperspectral imaging: characterization of a noninvasive, in vivo system for determining tissue perfusion," *Anal. Chem.* **74**(9), 2021–2028 (2002).
65. S. A. Shah, N. Bachrach, S. J. Spear, D. S. Letbetter, R. A. Stone, R. Dhir, J. W. Prichard, H. G. Brown, and W. A. LaFramboise, "Cutaneous wound analysis using hyperspectral imaging," *Biotechniques* **34**(2), 408–413 (2003).
66. K. J. Zuzak, M. T. Gladwin, R. O. Cannon, III, and I. W. Levin, "Imaging hemoglobin oxygen saturation in sickle cell disease patients using noninvasive visible reflectance hyperspectral techniques: effects of nitric oxide," *Am. J. Physiol. Heart Circ. Physiol.* **285**(3), H1183–H1189 (2003).
67. R. Gillies, J. E. Freeman, L. C. Cancio, D. Brand, M. Hopmeier, and J. R. Mansfield, "Systemic effects of shock and resuscitation monitored by visible hyperspectral imaging," *Diabetes Technol. Ther.* **5**(5), 847–855 (2003).
68. S. G. Kong, Y. R. Chen, I. Kim, and M. S. Kim, "Analysis of hyperspectral fluorescence images for poultry skin tumor inspection," *Appl. Opt.* **43**(4), 824–833 (2004).

69. B. Khoobehi, J. M. Beach, and H. Kawano, "Hyperspectral imaging for measurement of oxygen saturation in the optic nerve head," *Invest. Ophthalmol. Vis. Sci.* **45**(5), 1464–1472 (2004).
70. T. Vo-Dinh, D. L. Stokes, M. B. Wabuyele, and M. E. Martin, "A hyperspectral imaging system for in vivo optical diagnostics. Hyperspectral imaging basic principles, instrumental systems and applications of biomedical interest," *IEEE Eng. Med. Biol. Mag.* **23**(5), 4049 (2004).
71. A. J. Chaudhari, F. Darvas, J. R. Bading, R. A. Moats, P. S. Conti, D. J. Smith, S. R. Cherry, and R. M. Leahy, "Hyperspectral and multispectral bioluminescence optical tomography for small animal imaging," *Phys. Med. Biol.* **50**(23), 5421–5441 (2005).
72. B. S. Sorg, B. J. Moeller, O. Donovan, Y. Cao, and M. W. Dewhirst, "Hyperspectral imaging of hemoglobin saturation in tumor microvasculature and tumor hypoxia development," *J. Biomed. Opt.* **10**(4), 44004 (2005) [doi: 10.1117/1.2003369].
73. L. L. Randeberg, I. Baarstad, T. Løke, P. Kaspersen, and L. O. Svaasand, "Hyperspectral imaging of bruised skin," *Proc. SPIE* **6078**, 60780O (2006) [doi: 10.1117/12.646557].
74. L. L. Randeberg and J. Hernandez-Palacios, "Hyperspectral imaging of bruises in the SWIR spectral region," *Proc. SPIE* **8207**, 82070N (2012) [doi: 10.1117/12.909137].
75. L. L. Randeberg, "Hyperspectral characterization of tissue in the SWIR spectral range: a road to new insight?," *Proc. SPIE* **10873**, 108730Y (2019) [doi: 10.1117/12.2504297].
76. J. Burger and P. Geladi, "Hyperspectral NIR imaging for calibration and prediction: a comparison between image and spectrometer data for studying organic and biological samples," *Analyst* **131**(10), 1152–1160 (2006).
77. K. J. Zuzak, S. C. Naik, G. Alexandrakis, D. Hawkins, K. Behbehani, and E. H. Livingston, "Characterization of a near-infrared laparoscopic hyperspectral imaging system for minimally invasive surgery," *Anal. Chem.* **79**(12), 4709–4715 (2007).
78. H. Akbari, Y. Kosugi, K. Kojima, and N. Tanaka, "Detection and analysis of the intestinal ischemia using visible and invisible hyperspectral imaging," *IEEE Trans. Biomed. Eng.* **57**(8), 2011–2017 (2010).
79. H. Akbari, K. Uto, Y. Kosugi, K. Kojima, and N. Tanaka, "Cancer detection using infrared hyperspectral imaging," *Cancer Sci.* **102**(4), 852–857 (2011).
80. P. Usenik, M. Bürmen, A. Fidler, F. Pernuš, and B. Likar, "Near-infrared hyperspectral imaging of water evaporation dynamics for early detection of incipient caries," *J. Dent.* **42**(10), 1242–1247 (2014).
81. E. J. M. Baltussen, E. N. D. Kok, S. G. Brouwer de Koning, J. Sanders, A. G. J. Aalbers, N. F. M. Kok, G. L. Beets, C. C. Flohil, S. C. Bruin,

K. F. Kuhlmann, and H. J. Sterenborg, "Hyperspectral imaging for tissue classification, a way toward smart laparoscopic colorectal surgery," *J. Biomed. Opt.* **24**(1), 016002 (2019) [doi: 10.1117/1.JBO.24.1.016002].
82. Z. Liu, H. Wang, and Q. Li, "Tongue tumor detection in medical hyperspectral images," *Sensors* **12**(1), 162–174 (2012).
83. H. Fabelo, S. Ortega, R. Lazcano, D. Madroñal, G. M Callicó, E. Juárez, R. Salvador, D. Bulters, H. Bulstrode, A. Szolna, and J. F. Piñeiro, "An intraoperative visualization system using hyperspectral imaging to aid in brain tumor delineation," *Sensors* **18**(2), 430 (2018).
84. R. H. Wilson, K. P. Nadeau, F. B. Jaworski, R. Rowland, J. Q. M. Nguyen, C. Crouzet, R. B. Saager, B. Choi, B. J. Tromberg, and A. J. Durkin, "Quantitative short-wave infrared multispectral imaging of in vivo tissue optical properties," *J. Biomed. Opt.* **19**(8), 086011 (2014) [doi: 10.1117/1.JBO.19.8.086011].
85. D. C. Sordillo, L. A. Sordillo, P. P. Sordillo, L. Shi, and R. R. Alfano, "Short wavelength infrared optical windows for evaluation of benign and malignant tissues," *J. Biomed. Opt.* **22**(4), 045002 (2017) [doi: 10.1117/1.JBO.22.4.045002].
86. A. N. Bashkatov, E. A. Genina, and V. V. Tuchin, "Optical properties of skin, subcutaneous and muscle tissues: a review," *J. Innov. Opt. Health Sci.* **4**(1), 9–38 (2011).
87. M. Halicek, J. V. Little, X. Wang, M. Patel, C. C. Griffith, A. Y. Chen, and B. Fei, "Tumor margin classification of head and neck cancer using hyperspectral imaging and convolutional neural networks," *Proc. SPIE* **10576**, 1057605 (2018) [doi: 10.1117/12.2293167].
88. M. Halicek, H. Fabelo, S. Ortega, J. V. Little, X. Wang, A. Y. Chen, G. M. Callico, L. Myers, B. D. Sumer, and B. Fei, "Hyperspectral imaging for head and neck cancer detection: specular glare and variance of the tumor margin in surgical specimens," *J. Med. Imaging* **6**(3), 035004 (2019) [doi: 10.1117/1.JMI.6.3.035004].
89. C. R. Tracy, J. D. Terrell, R. P. Francis, E. F. Wehner, J. Smith, M. Litorja, D. L. Hawkins, M. S. Pearle, J. A. Cadeddu, and K. J. Zuzak, "Characterization of renal ischemia using DLP hyperspectral imaging: a pilot study comparing artery-only occlusion versus artery and vein occlusion," *J. Endourol.* **24**(3), 321–325 (2010).
90. F. Blanco, M. López-Mesas, M. Valiente, S. Serranti, G. Bonifazi, and J. Havel, "Hyperspectral imaging based method for fast characterization of kidney stone types," *J. Biomed. Opt.* **17**(7), 076027 (2012) [doi: 10.1117/1.JBO.17.7.076027].
91. D. Khanna, D. E. Furst, P. J. Clements, Y. Allanore, M. Baron, L. Czirjak, O. Distler, I. Foeldvari, M. Kuwana, M. Matucci-Cerinic, and M. Mayes, "Standardization of the modified Rodnan skin score for

use in clinical trials of systemic sclerosis," *J. Scleroderma Relat. Disord.* **2**(1), 11–18 (2017).
92. G. V. Baranoski, A. Dey, and T. F. Chen, "Assessing the sensitivity of human skin hyperspectral responses to increasing anemia severity levels," *J. Biomed. Opt.* **20**(9), 095002 (2015) [doi: 10.1117/1.JBO.20.9.095002].
93. P. Wang, Y. Cao, M. Yin, Y. Li, S. Lv, L. Huang, D. Zhang, Y. Luo, and J. Wu, "Full-field burn depth detection based on near-infrared hyperspectral imaging and ensemble regression," *Rev. Sci. Instrum.* **90**(6), 064103 (2019).
94. T. Du, D. K. Mishra, L. Shmuylovich, A. Yu, H. Hurbon, S. T. Wang, and M. Y. Berezin, "Hyperspectral imaging and characterization of allergic contact dermatitis in the short-wave infrared," *J. Biophotonics* **13**(9), e202000040 (2020).
95. Y.-M. Chen, K.-L. Lai, H.-H. Chen, W.-N. Huang, C.-T. Lin, W.-C. Chao, J.-P. Chen, Y.-W. Fu, H.-M. Chen, and P.-W. Lui, "Hyperspectral imaging for skin assessment in systemic sclerosis: a pilot study," *Rheumatology* **59**(11), 3201–3210 (2020).
96. L. L. Randeberg and J. Hernandez-Palacios, "Hyperspectral imaging of bruises in the SWIR spectral region," *Proc. SPIE* **8207**, 82070N (2012) [doi: 10.1117/12.909137].
97. H. Akbari, Y. Kosugi, K. Kojima, and N. Tanaka, "Hyperspectral imaging and diagnosis of intestinal ischemia," *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* **2008**, 1238–1241 (2008).
98. J. G. Rosas and M. Blanco, "A criterion for assessing homogeneity distribution in hyperspectral images. Part 2: application of homogeneity indices to solid pharmaceutical dosage forms," *J. Pharm. Biomed. Anal.* **70**, 691–699 (2012).
99. S. Kumar, C. Desmedt, D. Larsimont, C. Sotiriou, and E. Goormaghtigh, "Change in the microenvironment of breast cancer studied by FTIR imaging," *Analyst* **138**(14), 4058–4065 (2013).
100. D. Ravi, H. Fabelo, G. M. Callic, and G. Z. Yang, "Manifold embedding and semantic segmentation for intraoperative guidance with hyperspectral brain imaging," *IEEE Trans. Med. Imaging* **36**(9), 1845–1857 (2017).
101. M. F. Escoriza, J. M. VanBriesen, S. Stewart, J. Maier, and P. J. Treado, "Raman spectroscopy and chemical imaging for quantification of filtered waterborne bacteria," *J. Microbiol. Methods* **66**(1), 63–72 (2006).
102. T. L. Kammies, M. Manley, P. A. Gouws, and P. J. Williams, "Differentiation of foodborne bacteria using NIR hyperspectral imaging and multivariate data analysis," *Appl. Microbiol. Biotechnol.* **100**(21), 9305–9320 (2016).

103. M. Eady and B. Park, "Classification of Salmonella enterica serotypes with selective bands using visible/NIR hyperspectral microscope images," *J. Microsc.* **263**(1), 10–19 (2016).
104. H.-Y. N. Holman, R. Miles, Z. Hao, E. Wozei, L. M. Anderson, and H. Yang, "Real-time chemical imaging of bacterial activity in biofilms using open-channel microfluidics and synchrotron FTIR spectromicroscopy," *Anal. Chem.* **81**(20), 8564–8570 (2009).
105. Y.-Z. Feng, G. ElMasry, D.-W. Sun, A. G. M. Scannell, D. Walsh, and N. Morcy, "Near-infrared hyperspectral imaging and partial least squares regression for rapid and reagentless determination of Enterobacteriaceae on chicken fillets," *Food Chem*. **138**(2), 1829–1836 (2013).
106. B. K. Alsberg, T. Løke, and I. Baarstad, "PryJector: a device for in situ visualization of chemical and physical property distributions on surfaces using projection and hyperspectral imaging," *J. Forensic Sci.* **56**(4), 976–983 (2011).
107. K. Książek, M. Romaszewski, P. Głomb, B. Grabowski, and M. Cholewa, "Blood stain classification with hyperspectral imaging and deep neural networks," *Sensors* **20**(22), 6666 (2020).
108. C. Malegori, E. Alladio, P. Oliveri, C. Manis, M. Vincenti, P. Garofano, F. Barni, and A. Berti, "Identification of invisible biological traces in forensic evidences by hyperspectral NIR imaging combined with chemometrics," *Talanta* **215**, 120911 (2020).
109. L. L. Randeberg, A. M. Winnem, N. E. Langlois, E. L. P. Larsen, R. Haaverstad, B. Skallerud, O. A. Haugen, and L. O. Svaasand, "Skin changes following minor trauma," *Lasers Surg. Med.* **39**(5), 403–413 (2007).
110. J. Fisher and W. C. Welch, "Survey and analysis of fore-optics for hyperspectral imaging systems," *Proc. SPIE* **6206**, 62062R (2006) [doi: 10.1117/12.668526].
111. A. Bjorgan, L. A. Paluchowski, and L. L. Randeberg, "Combined 3D model acquisition and autofocus tracking system for hyperspectral line-scanning devices," *Proc. SPIE* **10870**, 108700D (2019) [doi: 10.1117/12.2505593].
112. X. Cao, T. Yue, X. Lin, S. Lin, X. Yuan, Q. Dai, L. Carin, and D. J. Brady, "Computational snapshot multispectral cameras: toward dynamic capture of the spectral world," *IEEE Signal Process. Mag.* **33**(5), 95–108 (2016).
113. W.-H. Su and D.-W. Sun, "Fourier transform infrared and Raman and hyperspectral imaging techniques for quality determinations of powdery foods: a review," *Compr. Rev. Food Sci. Food Saf.* **17**(1), 104–122 (2018).
114. M. Milanic and B. Majaron, "Three-dimensional Monte Carlo model of pulsed-laser treatment of cutaneous vascular lesions," *J. Biomed. Opt.* **16**(12), 128002 (2011) [doi: 10.1117/1.3659205].

115. H. Zhang, D. Salo, D. M. Kim, S. Komarov, Y.-C. Tai, and M. Y. Berezin, "Penetration depth of photons in biological tissues from hyperspectral imaging in shortwave infrared in transmission and reflection geometries," *J. Biomed. Opt.* **21**(12), 126006 (2016) [doi: 10.1117/1.JBO.21.12.126006].
116. M. Milanic, L. A. Paluchowski, and L. L. Randeberg, "Hyperspectral imaging for detection of arthritis: feasibility and prospects," *J. Biomed. Opt.* **20**(9), 096011 (2015) [doi: 10.1117/1.JBO.20.9.096011].
117. R. Gautam, S. Vanga, F. Ariese, and S. Umapathy, "Review of multidimensional data processing approaches for Raman and infrared spectroscopy," *EPJ Techn. Instrum.* **2**(1), 1–38 (2015).
118. P. Geladi, B. Sethson, J. Nyström, T. Lillhonga, T. Lestander, and J. Burger, "Chemometrics in spectroscopy: part 2. Examples," *Spectrochim. Acta Part B Atomic Spectrosc.* **59**(9), 1347–1357 (2004).
119. P. Geladi, "Chemometrics in spectroscopy. Part 1. Classical chemometrics," *Spectrochim. Acta Part B Atomic Spectrosc.* **58**(5), 767–782 (2003).
120. R. G. Brereton, "Biological and medical applications of chemometrics," Chapter 10 in *Applied Chemometrics for Scientists*, Wiley, Chichester, United Kingdom, pp. 287–318 (2007).
121. J. Gabrielsson, N.-O. Lindberg, and T. Lundstedt, "Multivariate methods in pharmaceutical applications," *J. Chemom.* **16**(3), 141–160 (2002).
122. A. El-Gindy and G. M. Hadad, "Chemometrics in pharmaceutical analysis: an introduction, review and future perspectives," *J. AOAC Int.* **95**(3), 609–623 (2012).
123. R. Kumar and V. Sharma, "Chemometrics in forensic science," *TrAC Trends Anal. Chem.* **105**, 191–201 (2018).
124. P. Geladi and E. Dåbakk, "An overview of chemometrics applications in near infrared spectrometry," *J. Near Infrared Spectrosc.* **3**(3), 119–132 (1995).
125. M. Manley, "Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis of biological materials," *Chem. Soc. Rev.* **43**(24), 8200–8214 (2014).
126. P. Lasch, "Spectral pre-processing for biomedical vibrational spectroscopy and microspectroscopic imaging," *Chemom. Intell. Lab. Syst.* **117**, 100–114 (2012).
127. A. Rácz, D. Bajusz, and K. Héberger, "Chemometrics in analytical chemistry," Chapter 9 in *Applied Chemoinformatics: Acheivements and Future Opportunities*, T. Eng and J. Gastinger, Eds., Wiley, Hoboken, pp. 471–499 (2018).

128. J. M. Prats-Montalbán, A. de Juan, and A. Ferrer, "Multivariate image analysis: a review with applications," *Chemom. Intell. Lab. Syst.* **107**(1), 1–23 (2011).
129. P. Oliveri, C. Malegori, and M. Casale, "Chemometrics: multivariate analysis of chemical data," Chapter 3 in *Chemical Analysis of Food: Techniques and Applications*, 2nd edition, Y. Pico, Ed., Academic Press, Elsevier, pp. 33–76 (2020).
130. R. Bro and A. K. Smilde, "Centering and scaling in component analysis," *J. Chemom.* **17**(1), 16–33, (2003).
131. O. E. de Noord, "The influence of data preprocessing on the robustness and parsimony of multivariate calibration models," *Chemom. Intell. Lab. Syst.* **23**(1), 65–70 (1994).
132. P. Oliveri, C. Malegori, R. Simonetti, and M. Casale, "The impact of signal pre-processing on the final interpretation of analytical outcomes—A tutorial," *Anal. Chim. Acta* **1058**, 9–17 (2019).
133. J. M. Amigo, I. Martí, and A. Gowen, "Hyperspectral imaging and chemometrics: a perfect combination for the analysis of food structure, composition and quality," in *Data Handling in Science and Technology*, F. Marini, Ed., Elsevier, Amsterdam, pp. 343–370 (2013).
134. A. Savitzky and M. J. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Anal. Chem.* **36**(8), 1627–1639 (1964).
135. Å. Rinnan, F. v. d. Berg, and S. B. Engelsen, "Review of the most common pre-processing techniques for near-infrared spectra," *TrAC Trends Anal. Chem.* **28**(10), 1201–1222 (2009).
136. A. Bjorgan, M. Milanic, and L. L. Randeberg, "Estimation of skin optical parameters for real-time hyperspectral imaging applications," *J. Biomed. Opt.* **19**(6), 066003 (2014) [doi: 10.1117/1.JBO.19.6.066003].
137. K. Esbensen and P. Geladi, "Principal component analysis: concept, geometrical interpretation, mathematical background, algorithms, history, practice," in *Comprehensive Chemometrics: Chemical and Biochemical Data Analysis*, S. Brown, R. Tauler, and B. Walczak Eds., Elsevier, Amsterdam, pp. A211–A226 (2009).
138. P. Geladi, H. Grahn, and M. Manley, "Data analysis and chemometrics for hyperspectral imaging," Chapter 3 in *Raman, Infrared, and Near-Infrared Chemical Imaging*, S. Šašić and Y. Ozaki, Eds., Wiley, Hoboken, pp. 93–107 (2010).
139. A. Biancolillo and F. Marini, "Chemometric methods for spectroscopy-based pharmaceutical analysis," *Front. Chem.* **6**, 576 (2018).
140. N. Mobaraki and J. M. Amigo, "HYPER-Tools: A graphical user-friendly interface for hyperspectral image analysis," *Chemom. Intell. Lab. Syst.* **172**, 174–187 (2018).

141. G. Sciutto, P. Oliveri, S. Prati, M. Quaranta, S. Lanteri, and R. Mazzeo, "Analysis of paint cross-sections: a combined multivariate approach for the interpretation of μATR-FTIR hyperspectral data arrays," *Anal. Bioanal. Chem.* **405**(2), 625–633 (2013).
142. E. Catelli, L. L. Randeberg, B. K. Alsberg, K. F. Gebremariam, and S. Bracci, "An explorative chemometric approach applied to hyperspectral images for the study of illuminated manuscripts," *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **177**, 69–78 (2017).
143. M. G. H. Omran, A. P. Engelbrecht, and A. Salman, "An overview of clustering methods," *Intell. Data Anal.* **11**(6), 583–605 (2007).
144. T. N. Tran, R. Wehrens, and L. M. C. Buydens, "Clustering multispectral images: a tutorial," *Chemom. Intell. Lab. Syst.* **77**(1–2), 3–17 (2005).
145. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, Springer, New York (2009).
146. J. M. Amigo, J. Cruz, M. Bautista, S. Maspoch, J. Coello, and M. Blanco, "Study of pharmaceutical samples by NIR chemical-image and multivariate analysis," *TrAC* **27**(8), 696–713 (2008).
147. A. Ghosh, N. S. Mishra, and S. Ghosh, "Fuzzy clustering algorithms for unsupervised change detection in remote sensing images," *Inf. Sci.* **181**(4), 699–715 (2011).
148. A. Golshan, H. Abdollahi, S. Beyramysoltan, M. Maeder, K. Neymeyr, R. Rajkó, M. Sawall, and R. Tauler, "A review of recent methods for the determination of ranges of feasible solutions resulting from soft modelling analyses of multivariate data," *Anal. Chim. Acta* **911**, 1–13 (2016)
149. M. Ghaffari, N. Omidikia, and C. Ruckebusch, "Essential spectral pixels for multivariate curve resolution of chemical images," *Anal. Chem.* **91**(17), 10943–10948 (2019).
150. W. Windig, "Two-way data analysis: Detection of purest variables," in *Comprehensive Chemometrics: Chemical and Biochemical Data Analysis*, S. D. Brown, R. Tauler, and B. Walczak, Eds., Elsevier, Amsterdam, pp. 275–307 (2009).
151. E. R. Malinowski, "Obtaining the key set of typical vectors by factor analysis and subsequent isolation of component spectra," *Anal. Chim. Acta* **134**, 129–137 (1982).
152. K. J. Schostack and E. R. Malinowski, "Preferred set selection by iterative key set factor analysis," *Chemom. Intell. Lab. Syst.* **6**(1), 21–29 (1989).
153. A. Bogomolov and M. McBrien, "Mutual peak matching in a series of HPLC–DAD mixture analyses," *Anal. Chim. Acta* **490**(1–2), 41–58 (2003).

154. E. Catelli, L. L. Randeberg, H. Strandberg, B. K. Alsberg, A. Maris, and L. Vikki, "Can hyperspectral imaging be used to map corrosion products on outdoor bronze sculptures?," *J. Spectr. Imaging* **7**, a10 (2018).
155. F. Cuesta Sánchez, J. Toft, B. Van den Bogaert, and D. L. Massart, "Orthogonal projection approach applied to peak purity assessment," *Anal. Chem.* **68**(1), 79–85 (1996).
156. W. Windig and J. Guilment, "Interactive self-modeling mixture analysis," *Anal. Chem.* **63**(14), 1425–1432 (1991).
157. W. Windig and D. A. Stephenson, "Self-modeling mixture analysis of second-derivative near-infrared spectral data using the SIMPLISMA approach," *Anal. Chem.* **64**(22), pp. 2735–2742 (1992).
158. Y. Batonneau, J. Laureyns, J. C. Merlin, and C. Brémard, "Self-modeling mixture analysis of Raman microspectrometric investigations of dust emitted by lead and zinc smelters," *Anal. Chim. Acta* **446**(1-2), 23–37 (2001).
159. S. Šašić, "An in-depth analysis of Raman and near-infrared chemical images of common pharmaceutical tablets," *Applied Spectroscopy* **61**(3), pp. 239-250 (2007)
160. S. Gourvénec, X. Capron, and D. L. Massart, "Genetic algorithms (GA) applied to the orthogonal projection approach (OPA) for variable selection," *Anal. Chim. Acta*, **519**(1), 11–21 (2004).
161. L. Duponchel, W. Elmi-Rayaleh, C. Ruckebusch, and J. P. Huvenne, "Multivariate curve resolution methods in imaging spectroscopy: influence of extraction methods and instrumental perturbations," *J. Chem. Inf. Comp. Sci.* **43**(6), 2057–2067 (2003).
162. X. Zhang and R. Tauler, "Application of multivariate curve resolution alternating least squares (MCR-ALS) to remote sensing hyperspectral imaging," *Anal. Chim Acta*, **762**, 25–38 (2013).
163. A. Nardecchia and L. Duponchel, "Randomised SIMPLISMA: Using a dictionary of initial estimates for spectral unmixing in the framework of chemical imaging," *Talanta* **217**, 121024 (2020).
164. J. Ferré, "Regression diagnostics," *in Comprehensive Chemometrics*, S. D. Brown, R. Tauler, and B. Walczak, Eds., Elsevier, Oxford, United Kingdom, pp. 33–89 (2009).
165. S. Wold, M. Sjöström, and L. Eriksson, "PLS-regression: a basic tool of chemometrics," *Chemom. Intell. Lab. Syst.* **58**(2), 109–130 (2001).
166. K. H. Esbensen, D. Guyot, F. Westad, and L. P. Houmoller, *Multivariate Data Analysis in Practice: An Introduction to Multivariate Data Analysis and Experimental Design*, Camo Process AS, Norway (2012).
167. C. P. Langlotz, B. Allen, B. J. Erickson, J. Kalpathy-Cramer, K. Bigelow, T. S. Cook, A. E. Flanders, M. P. Lungren, D. S. Mendelson, J. D. Rudie, and G. Wang, "A roadmap for foundational research on

artificial intelligence in medical imaging: from the 2018 NIH/RSNA/ACR/The Academy Workshop," *Radiology* **291**(3), 781–791 (2019).
168. A. Holzinger, P. Kieseberg, E. Weippl, and A. M. Tjoa, "Current advances, trends and challenges of machine learning and knowledge extraction: from machine learning to explainable AI," *Int. Cross-Domain Conf. for Machine Learning and Knowledge Extraction*, Springer, New York, 1–8 (2018).
169. G. Bologna and Y. Hayashi, "Characterization of symbolic rules embedded in deep DIMLP networks: a challenge to transparency of deep learning," *J. Artif. Intell. Soft Comput. Res.* **7**(4), 265–286 (2017).
170. W. Samek and K.-R. Müller, "Towards explainable artificial intelligence," in *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*, W. Samek, G. Montavon, A. Vedaldi, L. K. Hansen, and K.-R. Müller, Eds., Springer, New York, pp. 5–22 (2019).
171. V. Kolbjørnsrud, R. Amico, and R. J. Thomas, "Partnering with AI: how organizations can win over skeptical managers," *Strategy Leadersh.* **45**(1), 37–43 (2017).
172. A. Holzinger, C. Biemann, and C. S. Pattichis, "What do we need to build explainable AI systems for the medical domain?," *arXiv*:1712.09923, 1–28 (2017).
173. S. Trajanovski, C. Shan, P. J. C. Weijtmans, S. G. B. de Koning, and T. J. M. Ruers, "Tongue tumor detection in hyperspectral images using deep learning semantic segmentation," *IEEE Trans. Biomed. Eng.* **68**(4), 1330–1340 (2020).
174. N. Audebert, B. L. Saux, and S. Lefevre, "Deep learning for classification of hyperspectral data: a comparative review," *IEEE Geosci. Remote Sens. Mag.* **7**(2), 159–173 (2019).

Lise Lyngsnes Randeberg, Ph.D. is a professor of biomedical optics and photonics at the Norwegian University of Science and Technology (NTNU) in Trondheim, Norway. She has an M.Sc. in physics and mathematics (1999) and a Ph.D. in biomedical optics (2005) from NTNU. Her research interests include biomedical optics, light–tissue interactions, and applications of optical imaging and sensor technology for biomedical applications. This covers instrumentation, modeling, and data analysis for spectroscopy and hyperspectral imaging. She is a senior member of SPIE, a member of the Royal Norwegian Society of Science and Letters, and a member of the Norwegian Academy of Technical Sciences. In 2014 she received the Claude Rimington memorial award from the Norwegian Society of Photobiology and Photomedicine. In 2007 she won the Technoport's Young Innovator Award, and in 2006 she was awarded the American Society for Laser Surgery and Medicine Research Grant. She is a co-founder and co-owner of Picterus AS and a co-owner/part of the scientific team of Nordic Products AS. She has served as president of the Norwegian Society of Graduate Technical and Scientific Professionals from 2013–2021.

Emilio Catelli, Ph.D. is a postdoctoral fellow in the Department of Chemistry G. Ciamician of the University of Bologna, Italy. He obtained an M.Sc. in analytical chemistry from the University of Bologna and a Ph.D. in chemistry from the NTNU, Norway. His current research focuses on applying noninvasive and nondestructive spectroscopic techniques for investigation of cultural heritage artifacts. He specializes in developing chemometric approaches for processing VNIR, SWIR, and XRF hyperspectral imaging data.

Julio Hernández, M.Sc. is a senior research scientist at Norsk Elektro Optikk AS (NEO) in Oslo, Norway and manager of the Hyperspectral Applications Department at NEO's HySpex Division. He has a B.Sc. in physics from the National Autonomous University of Mexico (UNAM, Mexico City, Mexico) and an M.Sc. in nanotechnology from Chalmers University of Technology (Göteborg, Sweden). Julio has over 12 years of experience in the field of hyperspectral imaging, developing scientific-grade hyperspectral cameras and data acquisition systems for a variety of applications within remote sensing, defense, industry, and biomedical research. He is currently focused on developing customized solutions for end-users in academia and industry, and in promoting adoption of hyperspectral technologies in new markets.