

# Data acquisition and preprocessing

One of the most important tasks in preparation for data acquisition is to focus the object under investigation. The wavelength at which the detector has the highest sensitivity or quantum efficiency should initially be selected as the focal point. With known samples, it is of course also possible to focus on a defined wavelength (which is of particular interest for the evaluation). Particularly in the VNIR range, the possibility of binning (spectral and/or spatial summation of pixels on the detector) to increase intensity or minimize noise should also be examined. With higher intensities, in turn, the exposure time can be reduced and the recording frequency increased. In order to achieve a reliable evaluation of the hypercubes, necessary corrections such as bright and dark reference correction should already be carried out during data acquisition.

$$I_{\lambda} = \frac{(I_{\lambda, \text{measured}} - I_{\lambda, \text{dark reference}})}{I_{\lambda, \text{white reference}}}$$

*Typical spectroscopic corrections to avoid detector and system influences in measurements*

Bright field correction (or white reference or background correction) is carried out, for example, on optical PTFE, silicon wafers or similar spectrally inert materials. Dark field correction removes the detector offset from the data. Missing pixel detection is also important for MCT detectors. This is used to spectrally interpolate defective pixels (light and dark pixels, indicators) on the detector. Depending on the spectral resolution and spatial resolution, the resulting hypercubes can quickly reach sizes of several gigabytes. The way in which the hypercubes are viewed can vary: on the one hand, n images (with n = number of spectral bands/wavelengths) are available for data evaluation while on the other hand x\*y spectra are available.

Since information acquisition in the area is desired and an unbeatable advantage of HSI technology, the reduction of the data volume in the spectral direction is initially favorable. The actual information of the spectra, which is necessary for further evaluation, is advantageously separated from the noise by mathematical transformations, thus reducing the amount of data. The data pre-processing of the spectra therefore also works with the classic methods of spectroscopy, i.e:

- Baseline correction (e.g. linear, polynomial)
- Normalization (e.g. vector normalization, min/max normalization)
- Smoothing (e.g. Savitzky-Golay)
- Centering (e.g. mean value centering)
- Derivation(s)

For simpler questions, the user can get a first impression of which sample areas carry relevant information at which wavelengths by looking at individual wavelength images.

---

Revision #3

Created 8 August 2024 17:19:30 by Admin

Updated 8 August 2024 17:29:18 by Admin