DOI: 10.58895/ksp/1000150865 -3 erschienen in:
Forum Bildverarbeitung 2022 | Image Pocessing Forum 2022
DOI: 10.5445/KSP/1000150865 | https://www.ksp.kit.edu/site/books/m/10.58895/ksp/1000150865/

Areal multispectral sensor with variable choice of spatial and spectral resolution

Tobias Haist, Robin Hahn, and Stephan Reichelt

Universität Stuttgart, Institut für Technische Optik,
Pfaffenwaldring 9, 70569 Stuttgart, Germany

Abstract We present a newly developed method for snapshot multispectral imaging. The core idea is to use a diffractive optical element (DOE) in an intermediate image plane. The main advantages are the potentially cost effective implementation for different applications, e.g. for classification and the possibility to use different spatio-spectral samplings at different field positions. By appropriate choice of the DOE it is possible to chose the spectral and spatial sampling pattern. We also shortly address the issue of light efficiency for different approaches towards multispectral imaging.

Keywords Hyperspectral imaging, multispectral imaging, diffractive optics

1 Introduction

Spatially resolved spectral information can be fruitfully employed in a lot of applications ranging from food monitoring to the detection of air pollution. Most often, image sensors with so-called Bayer patterns are used which mimic the human visual system with three broad spectral channels, typically denoted as the “short” (blue), “middle” (green) and “long” (red) bands.

For some applications other or more spectral bands are advantageous. But one has to keep in mind that there is always a trade-off between spectral resolution, number of spectral channels, spatial resolution, light efficiency and measurement time. If high spectral and spatial resolution is desired, typically, the amount of light per spatio-spectral pixel element is low (for a given entrance pupil and luminance of the scene to be imaged). We can discriminate between "hyper-" and "multi-"spectral imaging based on the number of spectral channels. Typically, "hyper" is used for a lot of channels (e.g. more than 10). In the following we use the term "multi" if more than one channel is used. Therefore, even a standard RGB sensor is defined to be a "multi-spectral" sensor.

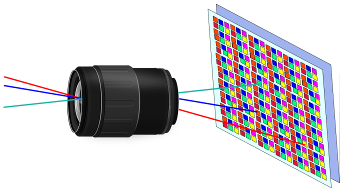
In general one can distinguish between snapshot and scanning sensors. Typically, for a small number of channels snapshot sensors are possible whereas for a large number of channels scanning approaches are applied (most often line-by-line imaging, so-called "push-broom imaging"). In this contribution we focus on snapshot imaging. An excellent overview and review is given by Hagen et al. in [1] and in the following we only will mention the main methods without going into detail about all possible sub-variants. Fig. 1 shows the basic sensing principles that are employed.

Most often, mosaics of absorption-based filters are used (as in the traditional Bayer pattern). This approach has a lot of advantages and is very cheap in mass production.

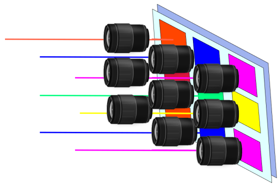
If more and narrower channels are desired, interference-based filters are employed [2]. Of course, the usable light per spatio-spectral sampling element is proportional to the spectral bandwidth and anti-proportional to the number of spectral channels if there is no spectral overlap (compare section 3). Therefore, for most applications one has to find a trade-off between spatio-spectral sampling and signal-to-noise ratio.

Anyway, disadvantages when using mosaics of dielectric filter have to be kept in mind. Homogeneous manufacturing of areas of such mosaics is complicated and expensive and the spectral response of a filter depends on the angle of incidence of the light (and neighbouring pixels). Therefore, image-sided telecentricity is advantageous. Anyway a thorough calibration of the sensor, ideally for every pixel, is necessary if really sensing with accurate spectral resolution is desired [2].

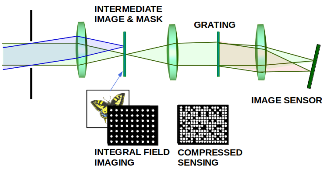
A variation of the standard mosaic approach is to use image replication. In this case for each of the replicated images an individual filter is employed. Filter manufacturing becomes easier but image replication has to be introduced. Most easily this can be realized macroscopicaly by just using several cameras side-by-side, each one equipped with one individual filter. However, for three-dimensional scenes there is a par-



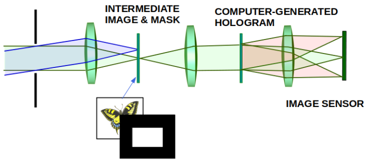
(a) Mosaic: absorption or interference



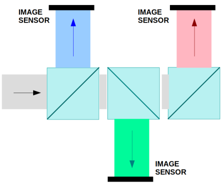
(b) Replication: separate pupils



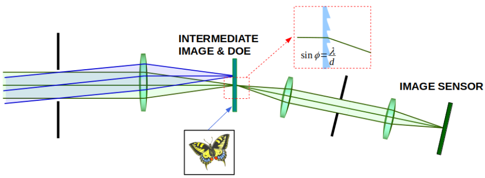
(c) Integral field imaging and compressed sensing



(d) CTIS



(e) Replication: spectral



(f) Diffraction-based

Figure 1: Basic sensing principles for snapshot multi-spectral imaging.

allax between the individual images that should be somehow corrected by image post-processing or otherwise leads to errors. The parallax error is proportional to the separation of the the entrance pupils of the individual image channels. Therefore, miniaturization is advantageous, leading e.g. to approaches like the one described by Hubold et al [3].

A classic alternative is to use the same entrance pupil and to split the image by dichroic beam splitters. This has been used a lot in commercial RGB color cameras because the light efficiency can be improved by that approach, of course at the cost of the need for multiple image sensors and their alignment.

If extensive post-processing is possible, so-called "computed tomography imaging spectroscopy" (CTIS) is an interesting option [4]. The light is diffracted by a computer-generated hologram in multiple diffraction orders that lead to separated copies of the scene on the image sensor. Each copy consists — again — of copies, one for each wavelength. On the sensor one obtains an overlap of all these wavelength separated copies and in the post-processing one tries to reconstruct the original multi-spectral information. Fast implementation is possible using neural networks [5].

The integral field approach uses spatial sampling with a pinhole array in combination with an imaging system with strong (lateral) chromatic aberration to obtain a spectrum for each of the sample points.

Now, if we open more pinholes the naive (and robust) sampling and dispersion approach will fail and we — again — will have overlap of information on the image sensor and some kind of reconstruction to obtain the spatially resolved spectral information is needed. Such approaches are typically denoted as "compressed sensing".

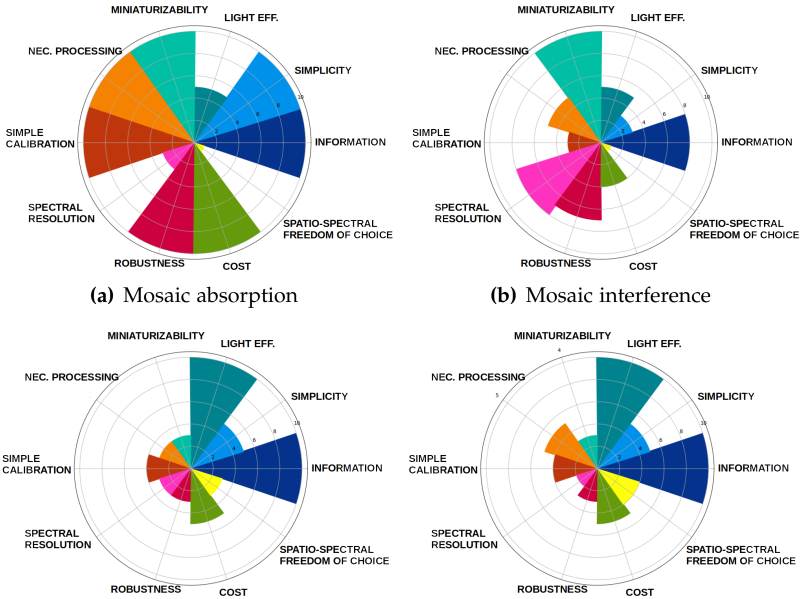
Fig. 2 and 2 show a qualitative comparison of the different sensing principles with respect to the key parameters of a snapshot hyperspectral sensor.

2 Diffraction-based multispectral sensor

One of the main disadvantages of mosaic-based multispectral imaging is the costly and difficult manufacturing of the mosaic filter. For high volume applications, of course, this is not an issue and such filters can be cheaply manufactured. But if specialized areas are to be realized, the initial development cost would be huge.

In Fig. 4 we show an alternative solution that uses diffraction instead of interference or absorption-based filters. It becomes possible to realize arbitrary spatio-spectral patterns by manufacturing a corresponding diffractive optical element. Such manufacturing is possible at rather manageable cost by several companies and universities.

The DOE is located in an intermediate image plane and deflects the light dependent on the wavelength. For understanding the working principle it is beneficial to first assume image-sided telecentricity of the first imaging stage and one large grating with constant grating pe-



(a) Mosaic absorption

(b) Mosaic interference

(c) Replication, separate pupils

(d) Replication, spectral

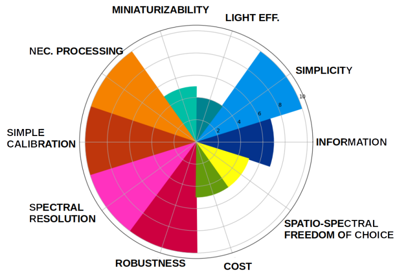
Figure 2: Qualitative comparison of different snapshot multi-spectral imaging approaches (large amplitudes are advantageous). Part 1

riod as the DOE. Due to the telecentricity, the chief rays will arrive at the same angle on the DOE and will be deflected according to their wavelengths.

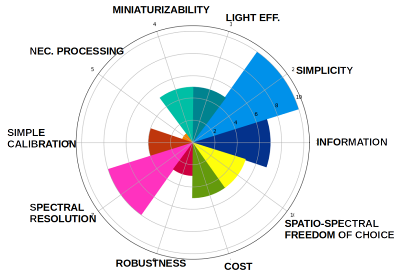
Different wavelengths then will hit the filter plane (actually the Fourier plane of the second imaging) at different locations and we can, therefore, let a certain spectrum pass the filter by using an appropriate iris. The rest of the second imaging system refocuses the light onto the monochrome (or color, if we want to combine with absorption-based filtering) image sensor.

By that approach we could make a sensor having a certain spectral response but only one channel.

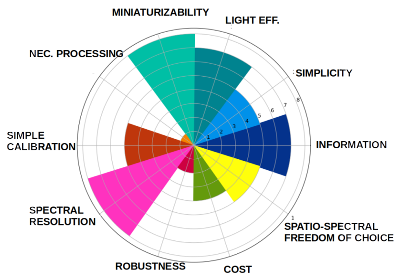
However, we can now replace the simple grating with a more complicated diffracting structure. E.g. we can use different micro-gratings



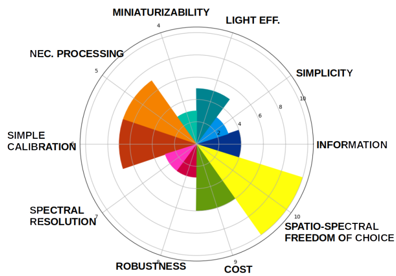
(a) Integral field imaging



(b) Compressed Sensing



(c) CTIS



(d) Diffraction-based

Figure 3: Qualitative comparison of different snapshot multi-spectral imaging approaches (large amplitudes are advantageous). Part 2

with different grating periods at different spatial locations in the intermediate image plane. We choose the periods of the gratings such that a certain wavelength will be deflected in the appropriate way so that it will pass the iris. Each "pixel" in the intermediate image will then consist of a micrograting and the grating period determines which wavelength will pass the iris.

Arbitrary spatio-spectral patterns can be realized by this kind of "grating-mosaic" and one can even realize complex spectra at one point by replacing a micro-grating with a more complex "computer-generated hologram".

Unfortunately, the spectral resolution is strongly coupled with the spatial resolution because the filter acts as a spectral filter *and* the aperture stop of the imaging at the same time. If we use a small hole as the

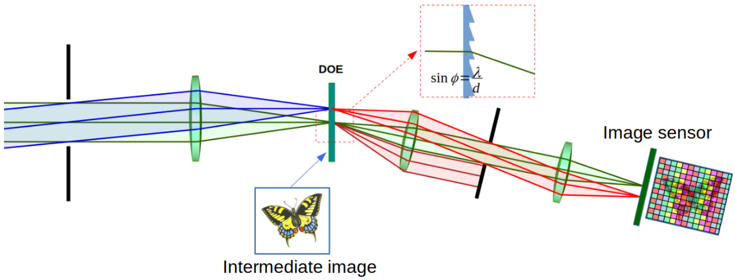


Figure 4: Principle of the diffraction based multi-spectral sensor.

iris, the spectral resolution $\Delta\lambda$ increases but the resolution according to Rayleigh Δr decreases. In [6] we derived the following uncertainty relation, which strongly depends on the minimal grating period d that can be manufactured (Δr is given in the intermediate image plane):

$$\Delta\lambda \cdot \Delta r \geq \lambda \cdot d \tag{1}$$

For a given minimum critical dimension of the DOE manufacturing d and a given size of image (and intermediate image) we will obtain a certain maximum number of resolution cells with a certain spectral bandwidth. This corresponds to the information that can be captured.

With the intermediate image size of $w \times h$ the information is given by

$$\Omega \approx \frac{w\,h}{\Delta r^2} \cdot \frac{\Lambda}{\Delta\lambda} \leq \frac{w\,h\Lambda\Delta\lambda}{\lambda^2\,d^2} \tag{2}$$

if the whole usable spectral range is denoted by Λ.

For an intermediate image with 20 mm $\times$ 20 mm, a minimum grating period of $d = 2\mu m$, a spectral bandwidth of 300 nm and a spectral resolution of 50 nm we obtain $\Omega \approx 5 \cdot 10^6$.

The less spectral channels we use, the larger the overall information that can be captured.

In Fig. 5 we show an example of a measurement with a 7 channel sensor where the DOE consists of stripes. Such an arrangement is es-

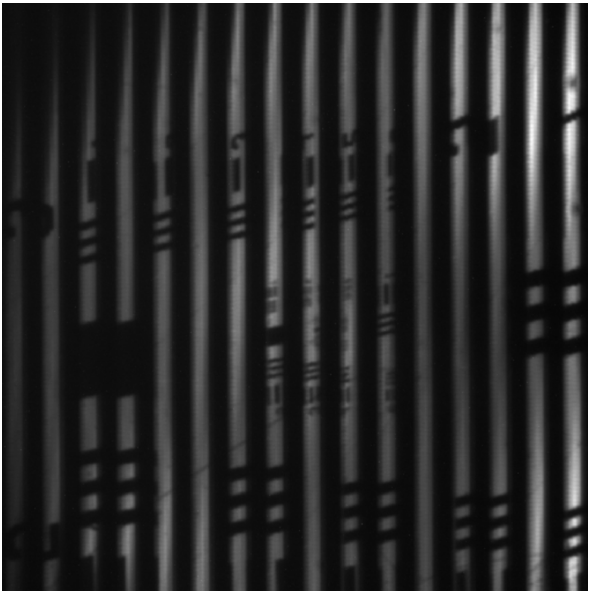


Figure 5: Signal on the image sensor during narrow-banded illumination of a USAF target. The USAF target was illuminated in transmission with a central wavelength of 632 nm. The half-width of the illumination spectrum was 10 nm.

pecially useful for detecting small shifts in wavelength. In the shown example spectral shift of 0.5 nm can be measured.

3 Light efficiency

Apart from the quite obvious parameters spatial and spectral resolution, the light efficiency is also very important. Good light efficiency allows one to use larger F-numbers or shorter exposure times at the same signal-to noise ratio.

We want to compare the different multispectral snapshot technologies according to the light efficiency. The baseline is a monochrome image sensor without any spectral channels.

The conventional absorption- or diffraction-based mosaic filter will be reduced the light efficiency by a factor

$$f_1 = \frac{\Delta\lambda}{\Lambda} \tag{3}$$

Beware that this is not the same than the number of spectral channels. It is advantageous most of the time to have a good light efficiency by using strongly overlapping filters. This is rarely done for commercial sensors but the standard in biology (compare e.g. the spectral responses of cones in the human eye). For classification purposes it is indeed often useful to employ overlapping channels and even simple processing can be used to classify based on spectral information.

In the human visual system, e.g. differences between the red and the green channel are "computed". The difference signal varies strongly with the spectrum of the input light if it lies in the overlapping region. Therefore, humans are extraordinary good in discerning green-yellowish colors. Obviously, object classification performance as well as light efficiency for ordinary scences is quite good.

The integral field imaging approach uses an amplitude mask in the intermediate image plane [7]. There is no spectral loss of light but, of course, the mask spatially filters and thereby eliminates a lot of photons. The separation of the individual pinholes should be at least N times larger than the diameter of the pinhole if we want to have N separated spectral channels. The associated loss is

$$f_2 = \frac{1}{N} \tag{4}$$

But again we could allow some kind of spectral overlap.

CTIS avoids the use of filtering at all. All incoming photons in principle (we neglect practical issues like the diffraction efficiency of the employed hologram) will arrive at the image sensor. However, it is not clear how to really compare with the filter-based pattern. Due to the overlapping of information a reconstruction step is necessary and at this stage noise might be amplified and artifacts might be introduced. Therefore, the really useful light efficiency is not 100% ($f = 1$). The overall noise is also increased because readout noise, quantization noise and fixed-pattern noise contributions will increase due to the effectively increased number of pixels that are exposed.

Compressed sensing in a snapshot approach lies somewhere between integral field imaging and CTIS. Again, information overlap will occur and, therefore, reconstruction is necessary. However with less information overlap compared to CTIS and less light reduction compared to integral field imaging.

The newly proposed diffraction based approach looses the light at the central iris. And the associated loss is simply again the same as with the integral field imaging approach f_2 if the individual channels are to be spectraly seperated. However, the approach might be worse because the F-number of the first image stage again is coupled with the spectral resolution: a large ray bundle would lead to bad spatial resolution. If one wants to achieve more spatial information with the same spectral behavior one has to increase the size of the intermediate image and as a result the whole setup becomes larger.

The conclusion is: All the approaches lead to more or less the same effective loss of light. The higher the spectral resolution (spectral half width of the channels) is chosen, the more loss is introduced.

One should carefully rethink if high spectral resolution anyway is necessary because overlapping channels are a good thing for a lot of applications.

4 Conclusion

The proposed sensor has the same light efficiency than other well known snapshot multispectral sensing principles. The main advantage is that one can freely chose spatial and spectral resolution at each position of the scene and that even more complex spectral responses can be easily realized using standard diffractive optics manufacturing.

However, spectral and spatial resolution are coupled by an uncertainty relation and also the F-number is coupled to the spectral resolution. In addition, an intermediate image is necessary. In practice this leads to increased space requirements for the sensor.

We thank the German ministry for education and research (BMBF) for financial support under the grant 13N15165 and Simon Amann for fruitful discussion on CTIS.

References

1. N. A. Hagen and M. W. Kudenov, “Review of snapshot spectral imaging technologies,” *Optical Engineering*, vol. 52, no. 9, p. 090901, 2013.
2. R. Hahn, F.-E. Hämmerling, T. Haist, D. Fleischle, O. Schwanke, O. Hauler, K. Rebner, M. Brecht, and W. Osten, “Detailed characterization of a mosaic based hyperspectral snapshot imager,” *Optical Engineering*, vol. 59, no. 12, p. 125102, 2020.
3. M. Hubold, R. Berlich, C. Gassner, R. Brüning, and R. Brunner, “Ultra-compact micro-optical system for multispectral imaging,” in *MOEMS and Miniaturized Systems XVII*, vol. 10545. SPIE, 2018, pp. 206–213.
4. T. Okamoto and I. Yamaguchi, “Simultaneous acquisition of spectral image information,” *Opt. Lett.*, vol. 16, no. 16, pp. 1277–1279, Aug 1991.
5. M. Zimmermann, S. Amann, M. Mel, T. Haist, and A. Gatto, “Deep learning-based hyperspectral image reconstruction from emulated and real computed tomography imaging spectrometer data,” *Optical Engineering*, vol. 61, no. 5, p. 053103, 2022.
6. R. Hahn, T. Haist, K. Michel, and W. Osten, “Diffraction-based hyperspectral snapshot imager,” *Optical Engineering*, vol. 61, no. 1, p. 015106, 2022.
7. R. Bacon, G. Adam, A. Baranne, G. Courtes, D. Dubet, J. Dubois, E. Emsellem, P. Ferruit, Y. Georgelin, G. Monnet *et al.*, “3d spectrography at high spatial resolution. i. concept and realization of the integral field spectrograph tiger.” *Astronomy and Astrophysics Supplement Series*, vol. 113, p. 347, 1995.