Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Review:

High-resolution spectral video acquisition*

Lin-sen CHEN†1, Tao YUE†‡1, Xun CAO1, Zhan MA1, David J. BRADY2

(1*School of Electronic Science and Engineering, Nanjing University, Nanjing 210023, China*)

(2*Department of Electrical & Computer Engineering, Duke University, Durham NC27708, USA*)

†E-mail: njucls@163.com; yuetao@nju.edu.cn

Received Feb. 8, 2017; Revision accepted July 12, 2017; Crosschecked Sept. 15, 2017

Abstract: Compared with conventional cameras, spectral imagers provide many more features in the spectral domain. They have been used in various fields such as material identification, remote sensing, precision agriculture, and surveillance. Traditional imaging spectrometers use generally scanning systems. They cannot meet the demands of dynamic scenarios. This limits the practical applications for spectral imaging. Recently, with the rapid development in computational photography theory and semiconductor techniques, spectral video acquisition has become feasible. This paper aims to offer a review of the state-of-the-art spectral imaging technologies, especially those capable of capturing spectral videos. Finally, we evaluate the performances of the existing spectral acquisition systems and discuss the trends for future work.

Key words: Multispectral/hyperspectral video acquisition; Snapshot; Under-sampling and reconstruction
https://doi.org/10.1631/FITEE.1700098 **CLC number:** TN919.8

1 Introduction

In mimicking human eyes, conventional RGB imaging devices perceive light with three types of micro-filters sensitive to three different parts of the light spectrum, respectively. Although trichromatic sensing suffices for human visual systems, spectral imaging can provide much more information about the captured scenes and objects. Since chemical elements have unique signatures in a spectral domain, multispectral/hyperspectral imaging technologies have a great application potential.

Traditional users of spectral imaging technology are in the fields of astronomy and remote sensing where the mapping of vegetation, minerals, water surfaces, and hazardous waste monitoring are of interest. In recent years, spectral images have been used increasingly in computer vision tasks, such as material discrimination (Du *et al.*, 2009), ophthalmology (Lawlor *et al.*, 2002), the study of combustion dynamics (Hunicz and Piernikarski, 2001), cellular dynamics (Kindzelskii *et al.*, 2000), surveillance (Harvey *et al.*, 2000), deciphering ancient scrolls (Mansfield, 2005), photography (Rørslett, 2004), medicine, agriculture, manufacturing, and forensics.

While spectral imaging is quite promising, the capture and processing of spectral data, especially high-dimensional spectral video data, faces significant challenges: First, as shown in Fig. 1, a spectral video holds one mega pixel in the spatial domain, hundreds of spectral channels, and tens of frames per second (FPS), and its capacity of within one second is about 10 GB. Second, traditional spectral acquisition systems commonly require specially-manufactured optical elements and complicated mechanical components, and cannot meet the demand of various low-cost and compact-size applications.

In this paper, we first review the traditional imaging spectrometers which are based on spatial scanning (Herrala *et al.*, 1994; Green *et al.*,

‡ Corresponding author

* Project supported by the National Natural Science Foundation of China (Nos. 61627804, 61371166, 61422107, 61571215, and 61671236) and the Natural Science Foundation of Jiangsu Province, China (Nos. BK20140610 and BK20160634)

ORCID: Lin-sen CHEN, http://orcid.org/0000-0002-1259-135X

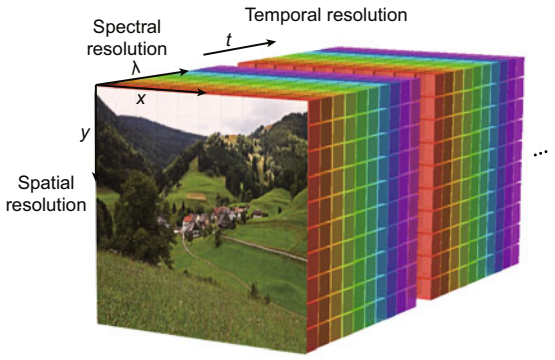

Fig. 1 Illustration of a high-dimensional spectral data cube. It spans two spatial dimensions (x, y), one spectral dimension (λ) and one temporal dimension (t)

1998) or spectral filtering (Morris *et al.*, 1994; Gat, 2000), trading the temporal or spatial resolution for spectral information. Then, we focus on several state-of-the-art computational imaging snapshot spectrometers: (1) the computed tomography imaging spectrometer (CTIS) (Descour and Dereniak, 1995; Descour *et al.*, 2001), (2) the coded aperture snapshot spectral imager (CASSI), from the early CASSI (Wagadarikar *et al.*, 2009) to its upgraded systems (Kittle *et al.*, 2010; Wu *et al.*, 2011; Lin *et al.*, 2014b), (3) the prism-mask multispectral video imaging system (PMVIS), including the low-spatial-resolution single-camera system (Cao *et al.*, 2011a), the high-spatial-resolution hybrid-camera system (Cao *et al.*, 2011b), and the content-adaptive hybrid-camera system (Ma *et al.*, 2014), and (4) some newly developed spectral acquisition systems based on the big-data science and semiconductor technologies, namely, the light field imaging spectrometer (LFIS) (Su *et al.*, 2015), the training-based spectrometer (Nguyen *et al.*, 2014; Oh *et al.*, 2016), and the colloidal quantum dot (CQD) spectrometer (Bao and Bawendi, 2015). In the end, we evaluate the performance of these existing spectral acquisition systems and discuss the trends for future work.

2 Traditional imaging spectrometers

A traditional spectrometer records the high-spectral-resolution information of a single point by dispersing a beam of light with optical dispersers, e.g., a prism or grating. To acquire spectral images, traditional spectral imaging devices use scanning systems, which trade the temporal information for high spectral or spatial resolution. Distinguished by their scanning systems, there are mainly two types of scanning-based spectral imaging devices, namely, the spatial scanning spectrometer and the spectral filtering spectrometer.

2.1 Spatial scanning spectrometer

As shown in Fig. 2, spatial scanning spectrometers can be divided into whisk-broom and push-broom spectrometers. Whisk-broom spectrometers, such as the airborne visible/infrared imaging spectrometer (AVIRIS) (Green *et al.*, 1998), record a single pixel spectrum each time, and the platform moves pixel by pixel until the entire plane is recorded. To improve the scanning efficiency, push-broom spectrometers record a spatial slit spectrum each time, such as in the hyperspectral digital imagery collection experiment (HYDICE) system (Mitchell, 1995), and the entire scene can be captured by moving the slit continuously.

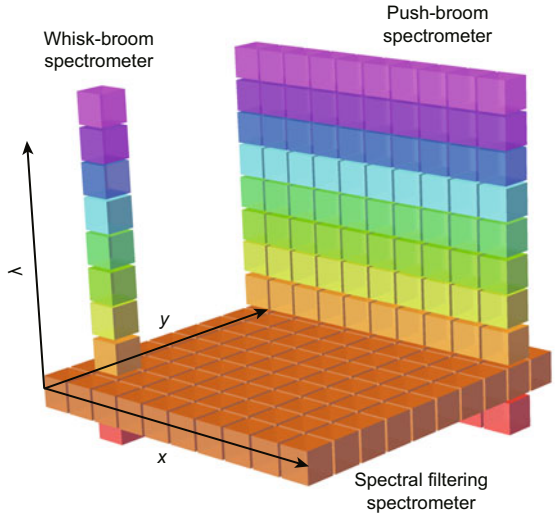

Fig. 2 Different scanning systems of traditional imaging spectrometers: whisk-broom spatial scanning, push-broom spatial scanning, and spectral filtering

2.2 Spectral filtering spectrometer

Spectral filtering spectrometers are based mainly on a set of different narrow bandpass color filters or electronically tunable filters, and the light intensity at different spectral channels can be recorded by switching the filters. Typical examples are: (1) the spatially varying color filters based camera in

Schechner and Nayar (2002), in which a spatially varying color filter is rigidly attached to a movable camera—as the camera moves, it senses each pixel in the scene multiple times, each time in a different spectral band; (2) the tunable filter-based imaging spectroscopy in Gat (2000), in which an electronically tunable filter is mounted in front of a monochrome camera, and spectral information is recorded by switching the pass-band of the filter; (3) the rotary filter-based system in Yamaguchi *et al.* (2006), in which several filters with different bandpass wavelengths are mounted into the rotating wheel, and the spectral images can be captured one after another by the monochrome camera placed behind the wheel.

To summarize, traditional imaging spectrometers produce additional spectral information by sampling continuous manifolds in a 3D data cube. To achieve high performance in terms of spectral accuracy, the voxels are sampled more evenly over the data cube to increase the information rate per captured frame. Thus, dynamic spectrum imaging becomes the bottleneck.

3 Computational imaging spectrometers

Since the scanning-based spectral imaging systems cannot capture dynamic scenes, technology for collecting high-dimensional spectral data cubes in a single snapshot appears to have great potential for various video-based applications, such as military target tracking, environmental pollution monitoring, and high-efficiency material classification.

The spectral data cube spans in three domains (x, y, and λ) with up to 10 GB capacity. Recording such an amount of data in a snapshot seems beyond the Nyquist-Shannon limit; however, it becomes practical by using the newly developed under-sampling and reconstruction technologies (Donoho, 2006; Candès and Wakin, 2008). The key feature determining the different reconstruction performances is the under-sampling strategy according to the optical configurations and statistical distribution properties of data (Candès *et al.*, 2006; Cao *et al.*, 2016). In this section, we will introduce several typical computational imaging spectrometers based on the under-sampling and reconstruction strategy.

3.1 Computed tomography imaging spectrometer

CTIS eliminates the need for scanning by acquiring full spectral information for all points within a 2D field of view during a single integration time. The original optical design and reconstruction algorithms of CTIS were introduced by Descour and Dereniak (1995). Some improvements were made later by Descour *et al.* (2001).

As shown in Fig. 3, a sequence of three transmission sinusoidal-phase gratings rotated in an increment of 60° can achieve dispersion in multiple directions into multiple orders. The dispersed images of the system's stop are interpreted as 2D projections of the 3D (x, y, λ) spectral cube. Then, the spectral cube is reconstructed using the maximum-likelihood and expectation-maximization algorithms (Shepp and Vardi, 1982) under the prior assumptions.

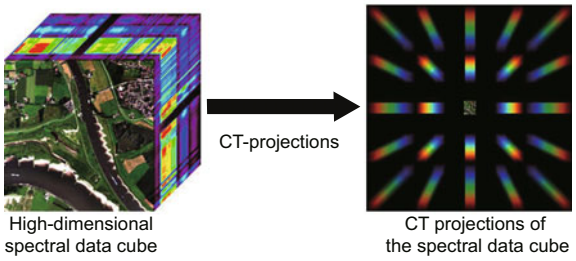

Fig. 3 Schematic of a computed tomography imaging spectrometer (CT: computed-tomography)

CTIS has the capacity for a spectrum snapshot, and the intensities at different wavelengths can be recorded directly without any color-filtering devices, which ensures the high throughput of light. However, computed-tomography reconstruction theory (Radon, 1917) requires continuous projection angles from 0° to 180°, while CTIS can record projection images only from a limited number of directions. These untapped projections leave a conical unsampled region in the 3D Fourier volume, which is commonly known as the 'missing cone problem' (Descour and Dereniak, 1995; Cao *et al.*, 2011a). The CTIS system faces an inevitably ill-posed limited problem caused by limited projection angles during the spectral imaging process. Filling the missing cone with additional computation was also discussed in Mooney *et al.* (1997).

In practice, the simultaneous capture of multiple diffraction orders reduces severely the transverse

spatial resolution in CTIS. The resolution would be better if only one diffraction order is captured and then the grating is rotated to obtain the diverse projections. Therefore, CTIS is not necessarily capable of acquiring spectral videos efficiently. Some more advanced video sampling techniques could be proposed.

3.2 Coded aperture snapshot spectral imager

Compressive sensing theory dictates that the high-dimensional spectral data cube can be reconstructed efficiently from far fewer measurements than those required by traditional linear scanning spectral systems.

Inspired by this, Brady *et al.* first introduced the compressive sensing theory to solve the underdetermined problem in spectral imaging, relying on the assumption that natural scenes are sparse inherently on some multi-scale basis (Wagadarikar *et al.*, 2009). To implement compressive spectral imaging, the aperture, which is used commonly as a light energy adjustor in conventional cameras, is designed specifically for coding the incoming rays. The imaging system is illustrated in Fig. 4 where light coming from the scene passes through the object lens first; then coupling with the coded aperture happens; next, the light is dispersed by the prism, and finally recorded by the sensor (CCD). Here, the coded aperture is considered to be binary, while the dispersive prism is considered linear. The imaging process can be expressed mathematically as (Arce *et al.*, 2014)

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{\theta} + \boldsymbol{\omega} = \boldsymbol{H}\boldsymbol{\psi}\boldsymbol{\theta} + \boldsymbol{\omega}, \tag{1}$$

where $\boldsymbol{A} = \boldsymbol{H}\boldsymbol{\psi}$ is the CASSI sensing matrix, $\boldsymbol{\theta}$ is a sparse representation of the spectral data cube on a 3D basis $\boldsymbol{\psi}$, $\boldsymbol{\omega}$ represents the noise of the system, and matrix $\boldsymbol{H}$ accounts for the effects of the coded aperture and the prism (Willett *et al.*, 2014). To reconstruct the 3D spectral data cube, given the set of measurements $\boldsymbol{y}$, the cost function (Golbabaee and Vandergheynst, 2012) should be minimized as follows:

$$\arg\min_{\boldsymbol{\theta}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{\theta}\| + \lambda \|\boldsymbol{\theta}\|_1, \tag{2}$$

where λ is a regularization constant. At this point, it should be emphasized that sensing matrix $\boldsymbol{A}$ plays a vitally important role in sensing to enable $\boldsymbol{\theta}$ to be as sparse as possible. In the following parts, we will present some upgrades to the CASSI system wherein more flexible sensing matrices are pursued.

Multi-frame CASSI: To improve the reconstruction accuracy, the proposed multi-frame CASSI, which was proposed by Kittle *et al.* (2010) provides more flexibility in strict adhering to the sparsity requirements needed for accurate estimation with compressive sensing. By taking multiple snapshots of the same scene with distinctly coded apertures, it is possible to select the number of measurements based on the resolution requirements while maintaining the snapshot advantages of the instrument in each unique shot. To realize a multiple snapshot, the static photomask is mounted on a piezostage so that, by spatially shifting the mask using the piezostage, different regions of the mask are exposed to the imaging scene. Furthermore, Wu *et al.* (2011) used a digital micromirror device (DMD) to implement the multiplexing pattern, substituting it for the static coded aperture. Since each pixel in DMD can be driven electrically to be turned on/off, it provides a virtually unlimited selection pool of multiplexing patterns.

Dual-coded compressive hyperspectral imaging: Whether single shot, multiple shot, or DMD-based, CASSI codes only the color spectrum in a spatially uniform manner. This sets a fundamental limit on the data quality that can be expected from the sparsity-constrained compressive reconstruction algorithms. Lin *et al.* (2014b) proposed a novel snapshot approach for compressive hyperspectral imaging, called 'dual-coded compressive hyperspectral imaging (DCSI)', wherein the spatial and spectral dimensions were coded, respectively, within a single snapshot.

Fig. 5 shows the schematic of the DCSI system, in which a two-arm system, including a spatial and a spectral modulation arm, is designed. In the spatial modulation arm, the objective lens is used to project the scene onto a DMD, which is

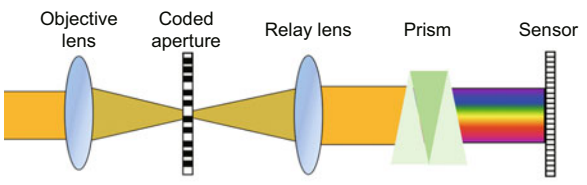

Fig. 4 Schematic of a coded aperture snapshot spectral imager

used as a high-resolution spatial light modulator. In the spectral modulation arm, a diffraction grating is brought in for spectrum dispersion, and a liquid crystal on silicon (LCOS) is added as the spectral modulator. During the exposure time, the spatial and spectral information can be modulated simultaneously and dynamically. Altogether, the reconstructed spectrum has been demonstrated to have a higher quality than the previous ones.

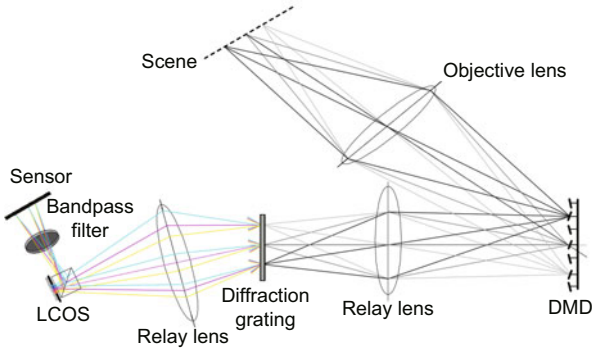

Fig. 5 Schematic of the dual-coded compressive hyperspectral imagin system (DMD: digital micromirror device) (Reprinted from Lin *et al.* (2014b), Copyright 2014, with permission from OSA)

To conclude, compressive sensing is a powerful sensing and reconstruction framework for recovering high-dimensional signals with only a few measurements. For spectral imaging, it offers a novel method for high-frame spectral video capture. When compared with traditional scanning spectrometers which rely on complicated mechanical scanning components, compressive spectral imaging systems are much more compact and flexible in various application fields (Arce *et al.*, 2014) and have lower cost. In addition, from the perspective of optical configurations, compressive sensing achieves much more sufficient light throughput than other spectrometers do, and can not only ensure the short exposure time for spectral video capture, but also result in high spectrum reconstruction accuracy.

However, limitations still exist in the coded aperture based imaging systems, whereby: (1) The reconstruction error is unavoidable due to the sparsity assumption for a natural scene (Willett *et al.*, 2014); (2) The computational complexities of the reconstruction algorithms, such as TwIST (Bioucas-Dias and Figueiredo, 2007), ADMM (Boyd *et al.*, 2011), GAP (Liao *et al.*, 2014), and HS-dictionary learning plus sparse-constraint computational reconstruction algorithms (Lin *et al.*, 2014a), are not satisfying, and the high-dimensional spectral data cannot be reconstructed in real time. This introduces obstacles for some time-critical applications.

3.3 Prism-mask multispectral video imaging system

In recent years, cameras have undergone rapid developments in spatial resolution, which are far beyond the displaying device's resolution and the perception of the human visual system. Meanwhile, spectral resolution has been the short board in many machine vision applications. Therefore, why not take the advantages of a camera's high spatial resolution for a higher spectral resolution?

Following this simple idea, Bodkin *et al.* (2009) proposed the Hyperpixel Array^{TM} Camera system, Gao *et al.* (2010) proposed an image mapping spectrometer, and Cao *et al.* (2011a) proposed PMVIS where the mask and micro lens array are used for spatial under-sampling. Combining the traditional spectroscopic methods, high-resolution cameras can capture the diffuse spectrum of sampling points; thus, spectral videos are captured allowing sacrifice in spatial resolution.

As shown in Fig. 6a, light emitted from the scene or object is down-sampled first by the uniform occlusion mask; the sampled light is then dispersed by the prism, and finally collected by the backward grayscale camera. Fig. 6b is an RGB image of the scene. Fig. 6c is the uniform occlusion mask used for spatial down-sampling, in which light passes only through the white rectangles, and the distance between the neighboring ones is well designed to avoid spectrum overlapping. Fig. 6d is the captured image, and the sampling rays have been diffused along the horizontal direction so that the intensities at different wavelengths can be recorded by different pixels. After the necessary calibrations and rectifications, the 3D spectral data cube can be extracted as shown in Fig. 6e.

Hybrid-camera system: Due to the sacrifice of spatial resolution for the additional spectral resolution, the image captured by PMVIS contains few spatial details. This may limit the ability to perform image analysis. Additionally, the higher the resolution in the spectral domain, the lower the resolution in the spatial domain. Assuming that the

RGB image retains most of the spatial edge structures of the high-dimensional image cube, a hybrid-camera system was proposed by Cao *et al.* (2011b). The schematic of the hybrid-camera system is illustrated in Fig. 7. Composed of an additional RGB camera and the PMVIS system, the hybrid-camera system can record two video streams simultaneously: an RGB video with a high spatial resolution, and a spectral video with a low spatial resolution. Following the registration of the two videos, the system propagates the spectral information into the RGB video stream so that a high-resolution spectral video is produced, in which both the proximity in the spatial domain and the color similarity are used to guide the propagation process, mathematically expressed as (Cao *et al.*, 2011b)

$$\boldsymbol{s}_{ij}=\sum_{c\in\{R,G,B\}}\frac{\sum\limits_{k\in\boldsymbol{\Omega}}G_{\sigma_r}(d_k^{\mathrm{RGB}})G_{\sigma_s}(d_k^{xy})\rho_k^c(w^c\otimes s_k)}{\sum\limits_{k\in\boldsymbol{\Omega}}G_{\sigma_r}(d_k^{\mathrm{RGB}})G_{\sigma_s}(d_k^{xy})},\tag{3}$$

where $\boldsymbol{s}_{ij}$ denotes the spectral vector of pixel (i,j), $k\in\boldsymbol{\Omega}$ indexes the pixels within the neighborhoods centered on (i,j), $G_{\sigma_s}(\cdot)$ represents the Gaussian operator with zero mean and variance σ, and d_k^{RGB} and d_k^{xy} denote the Euclidean distance between pixels (i,j) and k in the RGB space and (x,y) space, respectively, and factor ρ_k represents the ratio of a given color channel value at k to the corresponding value at (i,j).

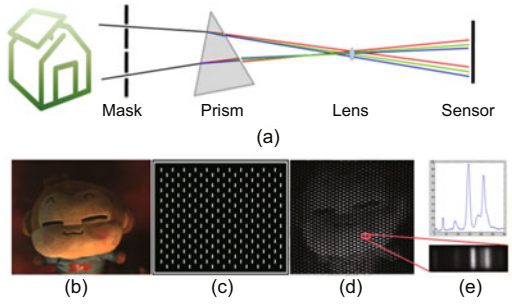

Fig. 6 Schematic of the prism-mask multispectral video imaging system: (a) light path; (b) an RGB image of the scene; (c) uniform occlusion mask; (d) intensities at different wavelengths of the diffused image; (e) a spectral signature of the 3D spectral data cube (References to color refer to the online version of this figure) (Reprinted from Cao *et al.* (2011a), Copyright 2011, with permission from IEEE)

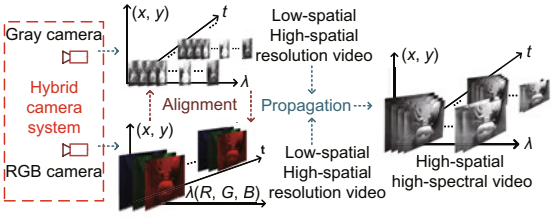

Fig. 7 Schematic of the hybrid-camera system (Reprinted from Cao *et al.* (2011b), Copyright 2011, with permission from IEEE)

Content-adaptive high-resolution hyperspectral video acquisition system: In addition, Ma *et al.* (2014) proposed a content-adaptive high-resolution hyperspectral video acquisition system to exploit fully the advantages of a hybrid-camera system. As illustrated in Fig. 8, compared with the hybrid-camera system, a spatial light modulator (SLM), where the sampling patterns are generated on-the-fly according to the scene content which is provided by the RGB camera, is used to replace the uniform occlusion mask. This leads to a more accurate and intelligent spectral video acquisition, which is more flexible for application in various application fields.

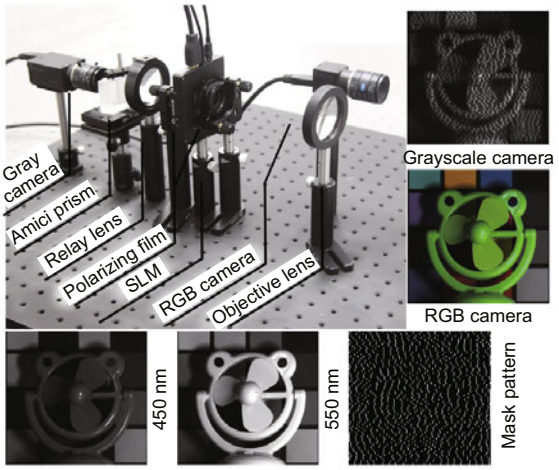

Fig. 8 The prototype and captured results for the content-adaptive high-resolution hyperspectral video acquisition system (Reprinted from Ma *et al.* (2014), Copyright 2014, with permission from OSA)

Unlike the traditional scanning spectrometers that trade either temporal or spatial resolution for additional spectral information, the prism-mask based hybrid-camera imaging system does not require such sacrifice in spatial details while maintaining high spectral accuracy. In contrast to coded aperture based imaging systems, these systems can be constructed with the low-cost off-the-shelf

optical components and are much simpler to calibrate in practice. Additionally, these systems can generate high-dimensional spectral videos in real time, and the effectiveness has been demonstrated with different computer vison applications including dynamic white balance adjustment and object tracking (Cao *et al.*, 2011b; Ma *et al.*, 2014).

3.4 More newly developed imaging systems

With the rapid development in big-data science and semiconductor device fabrication, some newly developed imaging systems have been emerging. In the following part, we introduce three typical systems in this group.

LFIS: As we know, a light field camera is able to simultaneously capture the spatial and angular information of the incoming rays emitted from the scene or object. A feasible spectral acquisition scheme was proposed by Zhou *et al.* (2010) from the perspective of a plenoptic function (Adelson and Bergen, 1991), in which the spectral information is coupled with the angular dimension information by placing a spectral filter at the aperture of the light field camera. LFIS can also capture the entire high-dimensional spectral data cube in a single shot, and the system maintains a compact size and weight. However, it still suffers from the tradeoff between spatial and spectral resolutions.

Training-based spectrometers: Since an RGB camera provides three measurements per pixel, it can be regarded as a spectral super-resolution problem to reconstruct the high-dimensional spectrum directly from the trichromatic measurements. The key to training-based methods is the mapping model from the captured RGB image and the high-dimensional spectral data; therefore, quite a few schemes have been presented, such as the spatio-spectral basis (Chakrabarti and Zickler, 2011), the metamer set (Morovic and Finlayson, 2006), and linear interpolation (Abed *et al.*, 2009) or non-linear interpolation (Nguyen *et al.*, 2014). Compared with other spectral imaging methods, the training-based ones have a much simpler and easy-to-use system. However, note that these single-image methods rely inevitably on strong assumptions and are extremely sensitive to the associated training data.

CQD spectrometer: Thanks to the developments in semiconductor device fabrication, a 2D absorptive filter array composed of CQDs has been coupled directly to the imaging sensor (Bao and Bawendi, 2015). Instead of measuring the spectra with gratings, prisms, or interference-based narrow-band filters, CQDs uses wide-band spectral filter arrays (Fig. 9a) to achieve high signal-to-noise ratio (SNR) measurements for reconstructing spectral images.

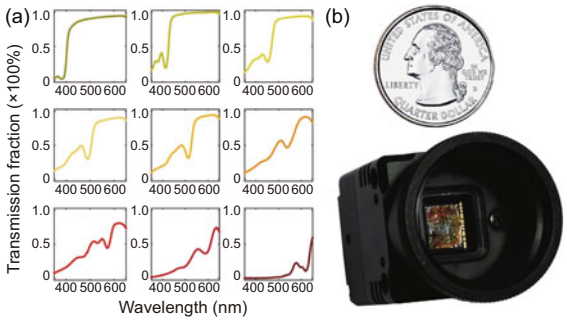

Fig. 9 The spectral transmission of different colloidal quantum dots (CQDs) (a) and the prototype of the CQD spectrometer (b) (Reprinted from Bao and Bawendi (2015), Copyright 2015, with permission from Macmillan Publishers Ltd.)

The imaging process of CQDs can be mathematically expressed as

$$I_i = \sum_{\lambda} \Phi(\lambda) T_i(\lambda),\ i = 1, 2, ..., n_{\mathrm{F}}, \qquad (4)$$

where I_i is the transmitted intensity ($i = 1, 2, ..., n_{\mathrm{F}}$ is the filter number), $T_i(\lambda)$ is the transmission spectral curve of the ith CQD filter, and $\Phi(\lambda)$ is the spectral intensity at wavelength λ of the incoming light. Because $T_i(\lambda)$ is predetermined by each CQD filter, the only unknown parameter is $\Phi(\lambda)$; thus, the linear regression algorithm can be performed for spectral reconstruction, by finding a spectrum $\hat{\Phi}(\lambda)$ to obtain the minimum mean square error (MSE).

The advantage of the CQD spectrometer is that it allows for a notable size reduction of the device. As shown in Fig. 9b, a quantum dot spectrometer has a size comparable to a US quarter, while it achieves a considerable performance. It has considerable potential in some size-critical applications. However, the spectral resolution of CQDs is quite limited and the maximum number of reconstructed spectral bands is equal to n_{F}, which is determined by the filter number. Furthermore, increase in the filter number requires a more sophisticated semiconductor technology. Meanwhile, too many filters lead inevitably to a decrease in the spatial resolution.

4 Evaluation and conclusions

Because spectral information can provide unique signatures of different chemical elements, approaches that enable efficient high-dimensional data acquisition appear to have enormous potential in various application fields.

As illustrated in Table 1, traditional imaging technologies use generally spatial scanning or spectral filtering systems to record the spectral information. They can capture the spectral images with both high spectral and spatial resolution, and the computation cost is quite low in that they record the spectra directly. However, since the scanning/filtering process is quite time-consuming, they cannot capture the spectral information of dynamic scenes. In addition, the embedded scanning/filtering mechanism is fairly complicated and expensive.

Fortunately, with the significant development in sampling theory and semiconductor device technology, several types of novel spectral imaging systems have emerged. Compared with traditional scanning-based spectrometers, computational spectral imaging systems have two significant advantages: (1) They can acquire spectral videos due to the breakthroughs in time-cost scanning to high-dimensional spectrum snapshots; (2) They can be constructed without specially manufactured optical components or complicated mechanical devices, endowing them with considerable potential in low-cost, size-critical applications.

On the other hand, the limitations of different computational spectrometers include: (1) CTIS: A conical unsampled region in the 3D Fourier volume is inevitable due to the untapped projections, and the construction has a high computation cost; (2) CASSI: The calibration complexity and computation cost are high; (3) PMVIS: The light throughput is poor on account of the occlusion mask; (4) LFIS and CQDs: The spatial and spectral resolutions in LFIS and the CQDs present a tradeoff, and the increase of either resolution leads to the decrease of the other; (5) The training-based spectrometers rely on strong assumptions and their accuracy is extremely sensitive to the training set.

To intuitively demonstrate the imaging quality, we use a public dataset (Yasuma *et al.*, 2010) to simulate the spectral images captured using different technologies. Noise and inference in the imaging process are ignored here. As shown in Fig. 10, it is obvious that the accuracy of the spectra captured by the training-based spectrometer is quite content-dependent, and the spectral details cannot be recovered by CASSI effectively. On the other hand, Fig. 11 shows the 2D image at 560 nm. The spatial details (Fig. 11b), which are captured by CASSI, are over-smoothed. The spatial resolutions of Figs. 11d and 11f are fairly low (which are respectively captured by LFIS and CQDs).

The central theme of this article is that, the under-sampling and reconstruction systems have demonstrated considerable potential in various applications, including both traditional remote sensing and computer vision tasks (segmentation, material discrimination, photography, surveillance, etc.), because of the remarkable efficiency in dynamic spectrum capture and simple low-cost system configurations. Meanwhile, several limitations exist in different spectral imaging systems, which must

Table 1 Comparisons of typical spectrometers

Spectrometer	Spectral resolution (nm)	Spatial resolution	Light throughput	Temporal resolution	Computation cost (min)	System complexity
Scanning	<1.0	10^6	Low	N/A	Low	High
CTIS	1.5–15.0	10^2	High	High	>10	High
CASSI	5.0–10.0	10^5	Medium	High	5–10	Medium
PMVIS	1.0–5.0	10^6	Low	Medium	Low	Medium
LFIS	Multi-band	10^4	Medium	Medium	Medium	Medium
Training	Multi-band	10^5	High	High	2	Low
CQDS	Multi-band	10^4	High	High	Low	Low

Scanning: traditional scanning based spectrometer (Herrala *et al.*, 1994; Green *et al.*, 1998; Gat, 2000); CTIS: computed-photography imaging spectrometer (Descour and Dereniak, 1995; Descour *et al.*, 2001); CASSI: coded aperture snapshot spectral imager (Wagadarikar *et al.*, 2009; Kittle *et al.*, 2010; Wu *et al.*, 2011; Lin *et al.*, 2014b); PMVIS: prism-mask multispectral video imaging system (Cao *et al.*, 2011a; 2011b; Ma *et al.*, 2014); LFIS: light field imaging spectrometer (Su *et al.*, 2015); Trainning: training-based spectrometer (Nguyen *et al.*, 2014; Oh *et al.*, 2016); CQDS: colloidal quantum dot spectrometer (Bao and Bawendi, 2015)

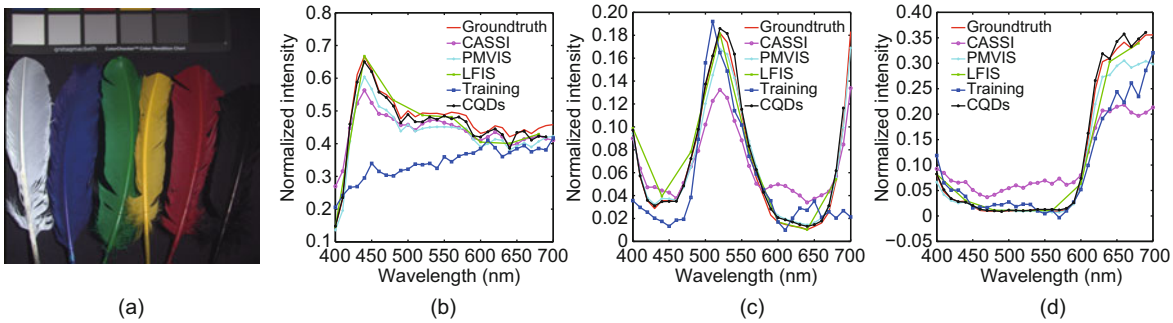

Fig. 10 The RGB reference image (a) and spectral signature comparisons for three areas: (b) white feather; (c) green feather; (d) red feather

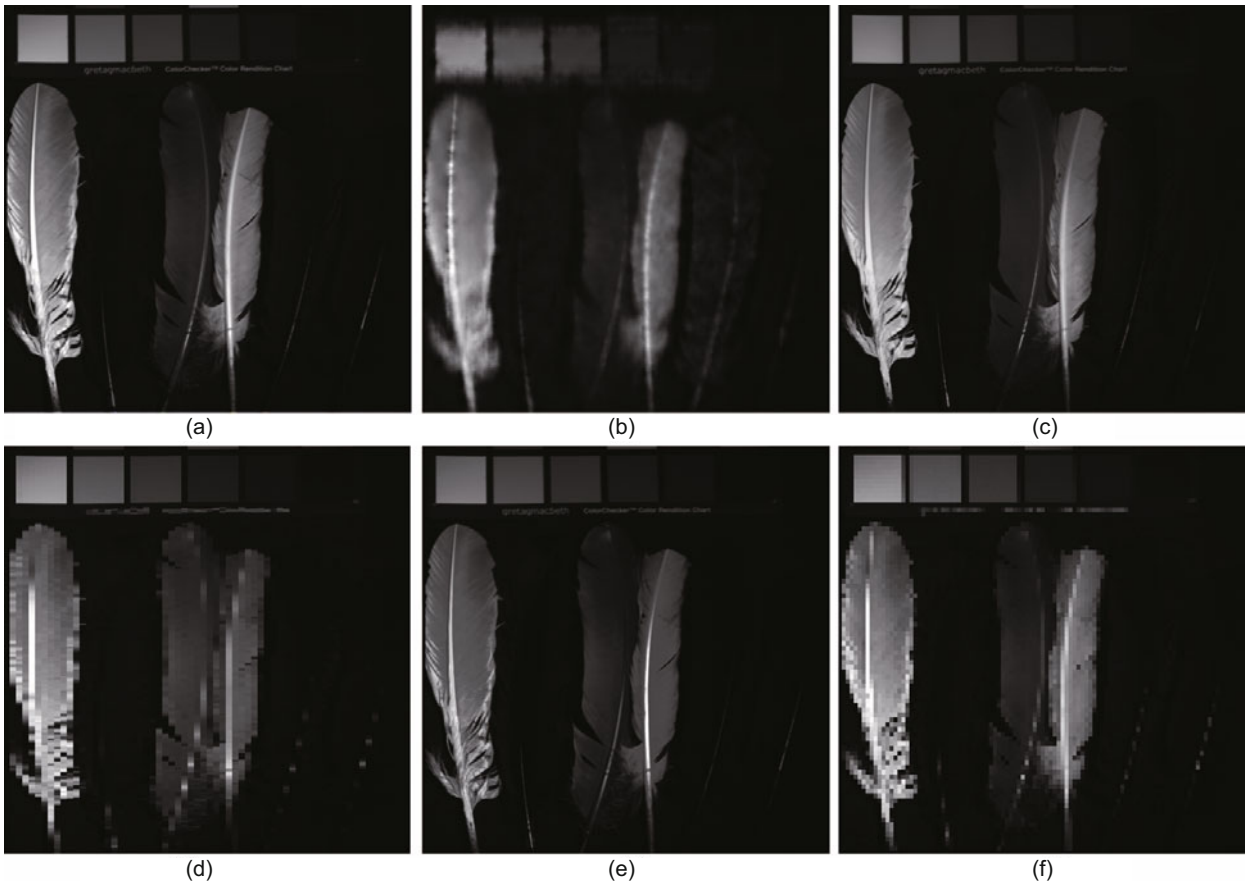

Fig. 11 Images at 560 nm which are captured by the traditional spectrometer (a), coded aperture snapshot spectral imager (b), prism-mask multispectral video imaging system (c), light field imaging spectrometer (d), training-based spectrometer (e), and the CQD spectrometer (f)

account for further combinations of optical principles, CS theory, machine learning algorithms, and semiconductor device fabrication technologies.

References

Abed, F.M., Amirshahi, S.H., Abed, M.R.M., 2009. Reconstruction of reflectance data using an interpolation technique. *J. Opt. Soc. Am. A*, **26**(3):613-624. https://doi.org/10.1364/JOSAA.26.000613

Adelson, E.H., Bergen, J.R., 1991. The plenoptic function and the elements of early vision. *In*: Landy, M.S., Movshon, J.A. (Eds.), Computational Models of Visual Processing. MIT Press, Cambridge, p.3-20.

Arce, G.R., Brady, D.J., Carin, L., *et al.*, 2014. Compressive coded aperture spectral imaging: an introduction.

IEEE Signal Process. Mag., **31**(1):105-115. https://doi.org/10.1109/MSP.2013.2278763

Bao, J., Bawendi, M.G., 2015. A colloidal quantum dot spectrometer. *Nature*, **523**(7558):67-70. https://doi.org/10.1038/nature14576

Bioucas-Dias, J.M., Figueiredo, M.A., 2007. A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration. *IEEE Trans. Imag. Process.*, **16**(12):2992-3004. https://doi.org/10.1109/TIP.2007.909319

Bodkin, A., Sheinis, A., Norton, A., *et al.*, 2009. Snapshot hyperspectral imaging: the hyperpixel array camera. *SPIE*, **7334**:73340H. https://doi.org/10.1117/12.818929

Boyd, S., Parikh, N., Chu, E., *et al.*, 2011. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.*, **3**(1):1-122. https://doi.org/10.1561/2200000016

Candès, E.J., Wakin, M.B., 2008. An introduction to compressive sampling. *IEEE Signal Process. Mag.*, **25**(2): 21-30. https://doi.org/10.1109/MSP.2007.914731

Candès, E.J., Romberg, J., Tao, T., 2006. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inform. Theory*, **52**(2):489-509. https://doi.org/10.1109/TIT.2005.862083

Cao, X., Du, H., Tong, X., *et al.*, 2011a. A prism-mask system for multispectral video acquisition. *IEEE Trans. Patt. Anal. Mach. Intell.*, **33**(12):2423-2435. https://doi.org/10.1109/TPAMI.2011.80

Cao, X., Tong, X., Dai, Q., *et al.*, 2011b. High resolution multispectral video capture with a hybrid camera system. IEEE Conf. on Computer Vision and Pattern Recognition, p.297-304. https://doi.org/10.1109/CVPR.2011.5995418

Cao, X., Yue, T., Lin, X., *et al.*, 2016. Computational snapshot multispectral cameras. *IEEE Signal Process. Mag.*, **33**(5):95-108. https://doi.org/10.1109/MSP.2016.2582378

Chakrabarti, A., Zickler, T., 2011. Statistics of real-world hyperspectral images. IEEE Conf. on Computer Vision and Pattern Recognition, p.193-200. https://doi.org/10.1109/CVPR.2011.5995660

Descour, M., Dereniak, E., 1995. Computed-tomography imaging spectrometer: experimental calibration and reconstruction results. *Appl. Opt.*, **34**(22):4817-4826. https://doi.org/10.1364/AO.34.004817

Descour, M., Volin, C.E., Ford, B.K., *et al.*, 2001. Snapshot hyperspectral imaging. *In*: Integrated Computational Imaging Systems. OSA Publishing, Washington, D.C., paper IWB4.

Donoho, D.L., 2006. Compressed sensing. *IEEE Trans. Inform. Theory*, **52**(4):1289-1306. https://doi.org/10.1109/TIT.2006.871582

Du, H., Tong, X., Cao, X., *et al.*, 2009. A prism-based system for multispectral video acquisition. IEEE 12th Int. Conf. on Computer Vision, p.175-182. https://doi.org/10.1109/ICCV.2009.5459162

Gao, L., Kester, R.T., Hagen, N., *et al.*, 2010. Snapshot image mapping spectrometer (IMS) with high sampling density for hyperspectral microscopy. *Opt. Expr.*, **18**(14):14330-14344. https://doi.org/10.1364Œ.18.014330

Gat, N., 2000. Imaging spectroscopy using tunable filters: a review. *SPIE*, **4056**:50-64. https://doi.org/10.1117/12.381686

Golbabaee, M., Vandergheynst, P., 2012. Compressed sensing of simultaneous low-rank and joint-sparse matrices. arXiv:1211.5058. http://arxiv.org/abs/1211.5058

Green, R.O., Eastwood, M.L., Sarture, C.M., *et al.*, 1998. Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS). *Remote Sens. Environ.*, **65**(3):227-248. https://doi.org/10.1016/S0034-4257(98)00064-9

Harvey, A.R., Beale, J.E., Greenaway, A.H., *et al.*, 2000. Technology options for imaging spectrometry. Int. Symp. on Optical Science and Technology, p.13-24. https://doi.org/10.1117/12.406592

Herrala, E., Okkonen, J.T., Hyvarinen, T.S., *et al.*, 1994. Imaging spectrometer for process industry applications. *SPIE*, **2248**:33-40. https://doi.org/10.1117/12.194344

Hunicz, J., Piernikarski, D., 2001. Investigation of combustion in a gasoline engine using spectrophotometric methods. *SPIE*, **4516**:307-314. https://doi.org/10.1117/12.435940

Kindzelskii, A.L., Yang, Z.Y., Nabel, G.J., *et al.*, 2000. Ebola virus secretory glycoprotein (sGP) diminishes FcγRIIIB-to-CR3 proximity on neutrophils. *J. Immun.*, **164**(2):953-958. https://doi.org/10.4049/jimmunol.164.2.953

Kittle, D., Choi, K., Wagadarikar, A., *et al.*, 2010. Multiframe image estimation for coded aperture snapshot spectral imagers. *Appl. Opt.*, **49**(36):6824-6833.

Lawlor, J., Fletcher-Holmes, D., Harvey, A., *et al.*, 2002. In vivo hyperspectral imaging of human retina and optic disc. *Invest. Ophthalmol. Vis. Sci.*, **43**(13):4350-4350. https://doi.org/10.1364/AO.49.006824

Liao, X., Li, H., Carin, L., 2014. Generalized alternating projection for weighted-$\ell_{2,1}$ minimization with applications to model-based compressive sensing. *SIAM J. Imag. Sci.*, **7**(2):797-823. https://doi.org/10.1137/130936658

Lin, X., Liu, Y., Wu, J., *et al.*, 2014a. Spatial-spectral encoded compressive hyperspectral imaging. *ACM Trans. Graph.*, **33**(6), Article 233. https://doi.org/10.1145/2661229.2661262

Lin, X., Wetzstein, G., Liu, Y., *et al.*, 2014b. Dual-coded compressive hyperspectral imaging. *Opt. Lett.*, **39**(7):2044-2047. https://doi.org/10.1364/OL.39.002044

Ma, C., Cao, X., Wu, R., *et al.*, 2014. Content-adaptive high-resolution hyperspectral video acquisition with a hybrid camera system. *Opt. Lett.*, **39**(4):937-940. https://doi.org/10.1364/OL.39.000937

Mansfield, C.L., 2005. Seeing into the Past. http://www.nasa.gov/vision/earth/technologies/scrolls.html

Mitchell, P.A., 1995. Hyperspectral digital imagery collection experiment (HYDICE). *SPIE*, **2587**:70-95. https://doi.org/10.1117/12.226807

Mooney, J.M., Vickers, V.E., An, M., *et al.*, 1997. High-throughput hyperspectral infrared camera. *J. Opt. Soc. Am. A*, **14**(11):2951-2961. https://doi.org/10.1364/JOSAA.14.002951

Morovic, P., Finlayson, G.D., 2006. Metamer-set-based approach to estimating surface reflectance from camera RGB. *J. Opt. Soc. Am. A*, **23**(8):1814-1822. https://doi.org/10.1364/JOSAA.23.001814

Morris, H.R., Hoyt, C.C., Treado, P.J., 1994. Imaging spectrometers for fluorescence and Raman microscopy: acousto-optic and liquid crystal tunable filters. *Appl. Spectr.*, **48**(7):857-866.

Nguyen, R.M., Prasad, D.K., Brown, M.S., 2014. Training-based spectral reconstruction from a single RGB image. European Conf. on Computer Vision, p.186-201. https://doi.org/10.1007/978-3-319-10584-0_13

Oh, W.S., Brown, M.S., Pollefeys, M., *et al.*, 2016. Do it yourself hyperspectral imaging with everyday digital cameras. IEEE Conf. on Computer Vision and Pattern Recognition, p.2461-2469. https://doi.org/10.1109/CVPR.2016.270

Radon, J., 1917. Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten. *Akad. Wiss.*, **69**:262-277 (in German).

Rørslett, B., 2004. All you ever wanted to know about digital UV and IR photography, but could not afford to ask. http://www.naturfotograf.com/UV_IR_rev00.html

Schechner, Y.Y., Nayar, S.K., 2002. Generalized mosaicing: wide field of view multispectral imaging. *IEEE Trans. Patt. Anal. Mach. Intell.*, **24**(10):1334-1348. https://doi.org/10.1109/TPAMI.2002.1039205

Shepp, L.A., Vardi, Y., 1982. Maximum likelihood reconstruction for emission tomography. *IEEE Trans. Med. Imag.*, **1**(2):113-122. https://doi.org/10.1109/TMI.1982.4307558

Su, L., Zhou, Z., Yuan, Y., *et al.*, 2015. A snapshot light field imaging spectrometer. *Opt.-Int. J. Light Electr. Opt.*, **126**(9):877-881. https://doi.org/10.1016/j.ijleo.2015.01.034

Wagadarikar, A.A., Pitsianis, N.P., Sun, X., *et al.*, 2009. Video rate spectral imaging using a coded aperture snapshot spectral imager. *Opt. Expr.*, **17**(8):6368-6388. https://doi.org/10.1364/OE.17.006368

Willett, R.M., Duarte, M.F., Davenport, M.A., *et al.*, 2014. Sparsity and structure in hyperspectral imaging: sensing, reconstruction, and target detection. *IEEE Signal Process. Mag.*, **31**(1):116-126. https://doi.org/10.1109/MSP.2013.2279507

Wu, Y., Mirza, I.O., Arce, G.R., *et al.*, 2011. Development of a digital-micromirror-device-based multishot snapshot spectral imaging system. *Opt. Lett.*, **36**(14):2692-2694. https://doi.org/10.1364/OL.36.002692

Yamaguchi, M., Haneishi, H., Fukuda, H., *et al.*, 2006. High-fidelity video and still-image communication based on spectral information: natural vision system and its applications. *SPIE*, **6062**:60620G. https://doi.org/10.1117/12.649454

Yasuma, F., Mitsunaga, T., Iso, D., *et al.*, 2010. Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum. *IEEE Trans. Imag. Process.*, **19**(9):2241-2253. https://doi.org/10.1109/TIP.2010.2046811

Zhou, Z., Yuan, Y., Bin, X.L., 2010. Light field imaging spectrometer: conceptual design and simulated performance. Frontiers in Optics/Laser Science XXVI, paper FThM3. https://doi.org/10.1364/FIO.2010.FThM3