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

High-resolution spectral video acquisition*

Lin-sen CHEN†[1], Tao YUE†[1], Xun CAO[1], Zhan MA[1], Jin-li SUO[2], David J. BRADY[3]

[1]*School of Electronic Science and Engineering, Nanjing University, Jiangsu 210023, China*

[2]*Department of Automation, Tsinghua University, Beijing 100084, China*

[3]*Department of Electrical & Computer Engineering, Duke University, Durham NC27708, USA*

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

Abstract: Compared with conventional cameras, spectral imagers provide many more features in the spectral domain, and they have been used in various fields such as material identification, remote sensing, precision agriculture, surveillance, etc. Traditional imaging spectrometers generally utilize scanning systems, which cannot meet the demands of dynamic scenarios, and 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 the ones capable of spectral video capturing. Finally, we evaluate the performance of these existing spectral acquisition systems and additionally discuss trends for future works.

Key words: multispectral/hyperspectral video acquisition, snapshot, under-sampling and reconstruction

doi:10.1631/FITEE.1000000 **Document code:** A **CLC number:**

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 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 increasingly been used in computer vision tasks, such as material discrimination (Hao Du and Lin, 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 (NASA's Jet Propulsion Laboratory,

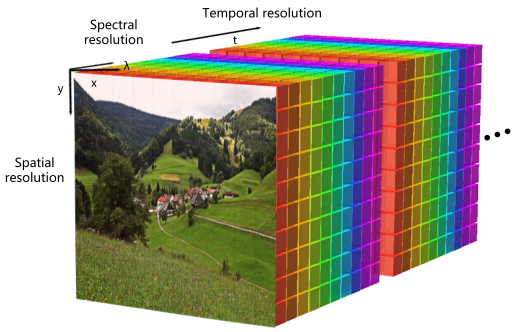


Fig. 1 Illustration of a high-dimensional spectral data cube. It spans two spatial dimensions (x,y), one spectral dimension (λ), and one temporal dimension (t).

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

ORCID: Linsen Chen, http://orcid.org/0000-0002-1259-135X

2005), photography (Rørslett, B, 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, the capacity of a spectral video within one second is about 10 gigabytes, which holds one mega pixel in the spatial domain, hundreds of spectral channels and tens of FPS (frames per second). Second, the traditional spectral acquisition systems commonly require specially-manufactured optical elements and complicated mechanical components, which cannot meet the demand of various low-cost, compact-size applications.

In this paper, we first review the traditional imaging spectrometers which are based on spatial scanning (Green *et al.*, 1998; Herrala *et al.*, 1994) or spectral filtering (Gat, 2000; Morris *et al.*, 1994), 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 *et al.*, 2001; Descour and Dereniak, 1995); (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; Wug Oh *et al.*, 2016) and the Colloidal Quantum Dot spectrometer (CQDs) (Bao and Bawendi, 2015). In the end, we evaluate the performance of these existing spectral acquisition systems and discuss trends for future works.

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 utilize 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.

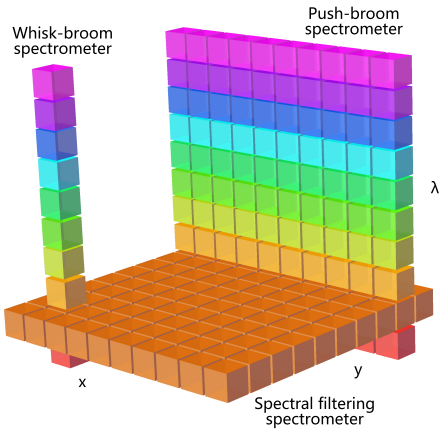


Fig. 2 The different scanning systems of traditional imaging spectrometers: (a) Whisk-broom spatial scanning; (b) Push-broom spatial scanning; (c) Spectral filtering.

2.1 Spatial Scanning Spectrometer

As shown in Fig. 2, spatial scanning spectrometers can be divided into the whisk-broom and the push-broom spectrometers. The whisk-broom spectrometers, such as AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) (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 HYDICE system (Hyperspectral Digital Imagery Collection Experiment) (Mitchell, 1995), and the entire scene can be captured by continuously moving the slit.

2.2 Spectral Filtering Spectrometer

The spectral filtering spectrometers are mainly based 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)

S. Nayar's spatially varying color filters based camera (Schechner and Nayar, 2002). In this system, 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) N. Gat's tunable filter-based imaging spectroscopy (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) M. Yamaguchi's rotary filter-based system (Yamaguchi *et al.*, 2006), where 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, the traditional imaging spectrometers produce additional spectral information by way of sampling continuous manifolds in the 3-D data cube. In order 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 are not capable of capturing dynamic scenes, technology for collecting the high-dimensional spectral data cube in a single snapshot appears to have great potential for various video-based applications, such as military targets tracking, environmental pollution monitoring, high-efficiency material classification, etc.

The spectral data cube spans in three domains (x, y, and λ) with up to 10 gigabytes of capacity. To record such an amount of data in a snapshot seems beyond the Nyquist-Shannon limit, but it becomes practical by utilizing 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 its 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)

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

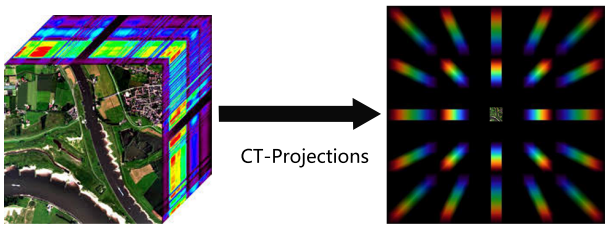


Fig. 3 A schematic diagram of a CTIS. Left: high-dimensional spectral data cube; Right: CT Projections of the spectral data cube.

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

Computed tomography imaging spectrometry has the capacity for a spectrum snapshot, and the intensity at different wavelengths can be directly recorded without any color filtering devices, which ensures the high light throughput. However, computed-tomography reconstruction theory (Radon, 1917) requires continuous projection angles from 0 to 180°, while CTIS can only record projection images from a limited number of directions. These untapped projections leave a conical unsampled region in the 3-D 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 the limited projection angles during

the spectral imaging process. Filling the missing cone with additional computation is also discussed in (Mooney *et al.*, 1997).

In practice, the simultaneous capture of multiple diffraction orders severely reduces the transverse spatial resolution in CTIS. The resolution is better if only one diffraction order is captured and then the grating is rotated to obtain diverse projections. Therefore, CTIS is not necessarily capable of efficient spectral video acquisition and some more advanced video sampling techniques could be proposed.

3.2 Coded Aperture Snapshot Spectral Imager (CASSI)

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

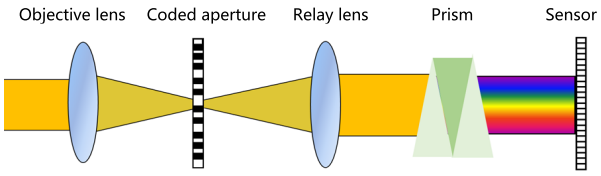


Fig. 4 The schematic diagram of a CASSI.

Inspired by this, D. Brady et al. first introduced compressive sensing theory to solve the underdetermined problem in spectral imaging, relying on the assumption that natural scenes are inherently sparse on some multi-scale basis (Wagadarikar *et al.*, 2009). To implement compressive spectral imaging, the aperture, which is commonly used as a light energy adjustor in conventional cameras, is specially designed 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 then finally recorded by the sensor (CCD). Here, the coded aperture is considered to be binary, and the dispersive prism is considered linear. The imaging process can be expressed mathematically as (Arce *et al.*, 2014):

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

Where $\boldsymbol{A} = \boldsymbol{H\psi}$ is the CASSI sensing matrix, $\boldsymbol{\theta}$ is a sparse representation of the spectral data cube on a three-dimensional basis $\boldsymbol{\psi}$, and $\boldsymbol{\omega}$ represents the noise of the system. The matrix $\boldsymbol{H}$ in equation (1) accounts for the effects of the coded aperture and the prism (Willett *et al.*, 2014). To reconstruct the 3-D spectral data cube, given the set of measurements $\boldsymbol{y}$, the cost function (Golbabaee and Vandergheynst, 2012) below should be minimized:

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

where λ is a regularization constant. At this point, it should be emphasized that the sensing matrix $\boldsymbol{A}$ plays a vitally important role in sensing to enable the $\boldsymbol{\theta}$ 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 Coded Aperture Snapshot Spectral Imaging: In order to improve the reconstruction accuracy, the multi-frame CASSI was proposed in 2010 (Kittle *et al.*, 2010), which provides more flexibility in strictly 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 still 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, Yuehao Wu et al. (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 electrically driven to be turned on and off, it provides a virtually unlimited selection pool of multiplexing patterns.

Dual-coded Compressive Hyperspectral Imaging: Whether a single shot, multiple shot, or DMD-based, CASSI only codes the color spectrum in a spatially uniform manner, which sets a fundamental limit on the data quality that can be expected from the compressive sparsity-constrained compressive reconstruction algorithms. In 2014, Lin et al. proposed a novel snapshot approach for compressive hyperspectral imaging called dual-coded compressive hyperspectral imaging (DCSI) (Lin *et al.*, 2014b),

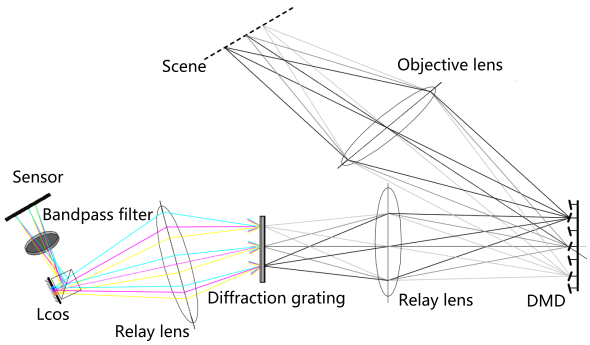


Fig. 5 The schematic drawing of the DCSI system (Lin *et al.*, 2014b).

wherein both spatial and spectral dimensions were coded, respectively, within a single snapshot.

Figure 5 shows the schematic drawing of the DCSI system, in which a two-arm system, including a spatial modulation arm 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 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, both the spatial and spectral information can be simultaneously modulated dynamically. Altogether, the reconstructed spectrum has been demonstrated to have a higher quality than the previous ones.

To conclude, compressive sensing is a powerful sensing and reconstruction framework for recovering high-dimensional signals with only a few measurements, and 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, low-cost, and flexible for various application fields (Arce *et al.*, 2014). In addition, from the perspective of optical configurations, it achieves much more sufficient light throughput than other spectrometers do, which can not only ensure the short exposure time for spectral video capture, but also results 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); and (2), the computational complexity 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, which introduces obstacles for some time-critical applications.

3.3 Prism-mask Multispectral Video Imaging System (PMVIS)

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

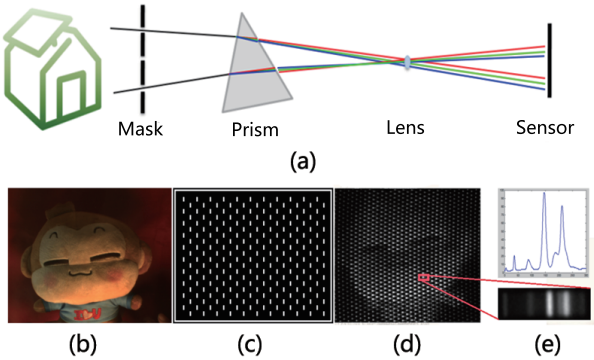


Fig. 6 The schematic diagram of PMVIS (Cao *et al.*, 2011a).

Following this simple idea, Andrew Bodkin proposed the Hyperpixel Array^{TM} Camera system in 2009 (Bodkin *et al.*, 2009); Liang Gao proposed an Image Mapping Spectrometer in 2010 (Gao *et al.*, 2010), and Xun Cao proposed the prism-mask multispectral video imaging system in 2011 (Cao *et al.*, 2011a), wherein the mask and micro lens array were utilized 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 the sacrifice in spatial resolution.

As shown in Fig. 6(a), light emitted from the

scene or object is first down-sampled by the uniform occlusion mask, then the sampled light is dispersed by the prism, and finally collected by the backward grayscale camera. Figure 6(b) is an RGB image of the scene. Figure 6(c) 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. Figure 6(d) is the captured image, and the sampling rays have been diffused along the horizontal direction so that the intensity at different wavelengths can be recorded by different pixels. After the necessary calibrations and rectifications, the 3-D spectral data cube can be extracted as is shown in Fig. 6(e).

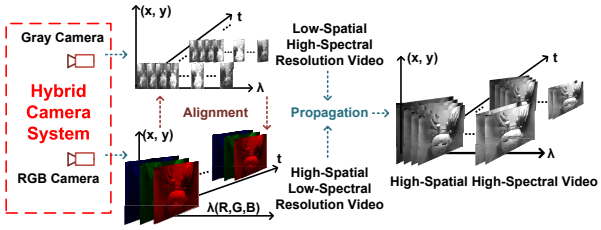


Fig. 7 The schematic diagram of the hybrid camera system (Cao *et al.*, 2011b).

Hybrid Camera System: Due to the sacrifice of spatial resolution for additional spectral resolution, the image captured by PMVIS contains few spatial details, which may limit the ability to perform image analysis. Additionally, the higher the resolution is in the spectral domain, the lower the resolution will be in the spatial domain. Under the assumption that the RGB image retains most of the spatial edge structures of the high-dimensional image cube, a hybrid camera system was proposed by Xun Cao in 2012 (Cao *et al.*, 2011b). The schematic diagram 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 simultaneously record two video streams: an RGB video with high spatial resolution, and a spectral video with low spatial resolution. Following 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):

$$s_{ij} = \sum_{c\in r,g,b} \frac{\sum_{k\in\Omega} G_{\sigma_r}(d_k^{RGB}) G_{\sigma_s}(d_k^{xy}) \rho_k^c (w^c \bigotimes s_k)}{\sum_{k\in\Omega} G_{\sigma_r}(d_k^{RGB}) G_{\sigma_s}(d_k^{xy})}, \tag{3}$$

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

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

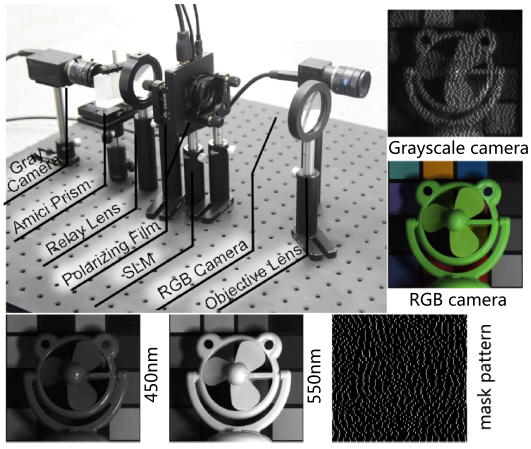


Fig. 8 The prototype and captured results for the content-adaptive high-resolution hyperspectral video acquisition system (Ma *et al.*, 2014).

Unlike the traditional scanning spectrometers that trade off either temporal or spatial resolution

for additional spectral information, the prism-mask-based hybrid-camera Imaging systems require no such sacrifices in spatial detail 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. What's more, 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 among this group.

Light Field Imaging Spectrometer (LFIS): As we know, a light field camera is able of simultaneously capturing 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. (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. The light field imaging spectrometer 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 resolution.

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 by (Chakrabarti and Zickler, 2011), the metamer set (Morovic and Finlayson, 2006), and linear (Abed *et al.*, 2009) or non-linear (Nguyen *et al.*, 2014) interpolation, and so on. Compared with other spectral imaging methods, the training-based ones have a much simpler and easy-to-use system. However, what is notable is that these single image methods inevitably rely on strong assumptions and are extremely sensitive to the associated training data.

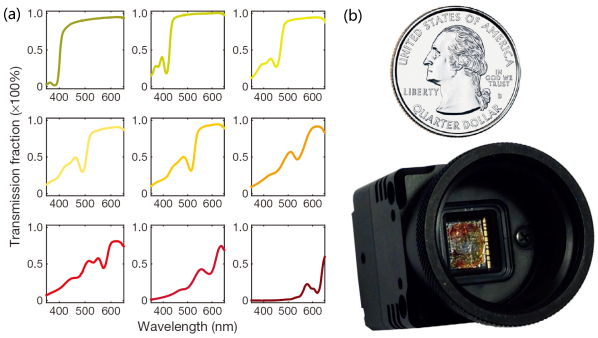


Fig. 9 (a) The spectral transmission of different colloidal quantum dots. (b) The prototype of the colloidal quantum dot spectrometer (Bao and Bawendi, 2015).

Colloidal Quantum Dot Spectrometer (CQDs): Thanks to the developments in semiconductor device fabrication, a kind of two-dimensional absorptive filter array composed of colloidal quantum dots (CQDs) has been directly coupled to the imaging sensor (Bao and Bawendi, 2015). Instead of measuring the spectra with gratings, prisms or interference-based narrowband filters, CQDs utilizes wide-band spectral filter arrays (as shown in Fig. 9(a)) to achieve high signal-to-noise ratio (SNR) measurements for reconstructing spectral images.

The imaging process of the CQDs can be mathematically expressed as:

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

Where I_i is the transmitted intensity ($i = 1, 2, ..., n_F$ is the filter number), $T_i(\lambda)$ is the transmission spectral curve of the CQD filters i, 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 is $\Phi(\lambda)$, so the linear regression algorithm can be performed for spectral reconstruction, by finding a spectrum $\hat{\Phi(\lambda)}$ to obtain the minimum MSE (mean square error).

The advantage of the colloidal quantum dot spectrometer is that it allows for a notable size re-

Table 1 Comparisons of typical spectrometers.

	Scanning	CTIS	CASSI	PMVIS	LFIS	Training	CQDs
spectral resolution	<1nm	1.5-15nm	5-10nm	1-5nm	multi-bands	multi-bands	multi-bands
spatial resolution	10^6	10^2	10^5	10^6	10^4	10^5	10^4
light throughput	low	high	medium	low	medium	high	high
temporal resolution	N/A	high	high	medium	medium	high	high
computation cost	low	>10min	5-10min	low	medium	2min	low
system Complexity	high	high	medium	medium	medium	low	low

* Scanning: the traditional scanning based spectrometers (Green *et al.*, 1998; Herrala *et al.*, 1994; Gat, 2000); CTIS: computed-photography imaging spectrometer (Descour *et al.*, 2001; Descour and Dereniak, 1995); 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 mutispectral video imaging system (Cao *et al.*, 2011a,b; Ma *et al.*, 2014); LFIS: light field imaging spectrometer (Su *et al.*, 2015); Trainning: training-based spectrometers (Nguyen *et al.*, 2014; Wug Oh *et al.*, 2016); CQDs: Colloidal Quantum Dot Spectrometer (Bao and Bawendi, 2015).

duction of the device. As is shown in Fig 9(b), a quantum dot spectrometer has a size comparable to a US quarter, while it achieves a considerable performance, and it has considerable potential in some size-critical applications. However, the spectral resolution of CQDs is quite limited and the maximum reconstructed spectral bands are equal to n_F, which is determined by the filter number. Furthermore, increases in the filter number require more sophisticated semiconductor technology. Meanwhile, too many filters inevitably lead to a decrease in the spatial resolution.

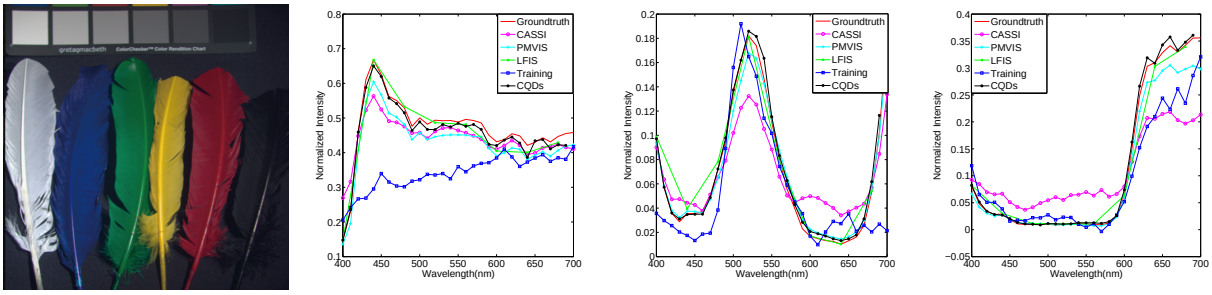
4 Evaluation and conclusion

Because the 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, the traditional imaging technologies generally utilize 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 developments in sampling theory and semiconductor device technology, several types of novel spectral imaging systems have emerged. When compared with the traditional scanning-based spectrometers, the computational spectral imaging systems have two significant advantages: (1) they are capable of spectral video acquisition due to the breakthroughs in time-cost scanning to high-dimensional spectrum snapshots; and (2), they can be constructed with no specially manufactured optical components and no complicated mechanical devices, which endows them with considerable potential in low-cost, size-critical applications.

On the other hand, the limitations of different computational spectrometers can be concluded to be: (1) CTIS: A conical unsampled region in the 3-D Fourier volume is inevitable due to the untapped projections, and the construction has a computation cost. (2) CASSI: The calibration complexity and computation cost is high. (3) PMVIS: the light throughput is poor on account of the occlusion mask.



(a) RGB image (b) (c) (d)

Fig. 10 (a) The RGB reference image. Spectral signature comparison for three areas: (b) white feather, (c) green feather, and (d) red feather.

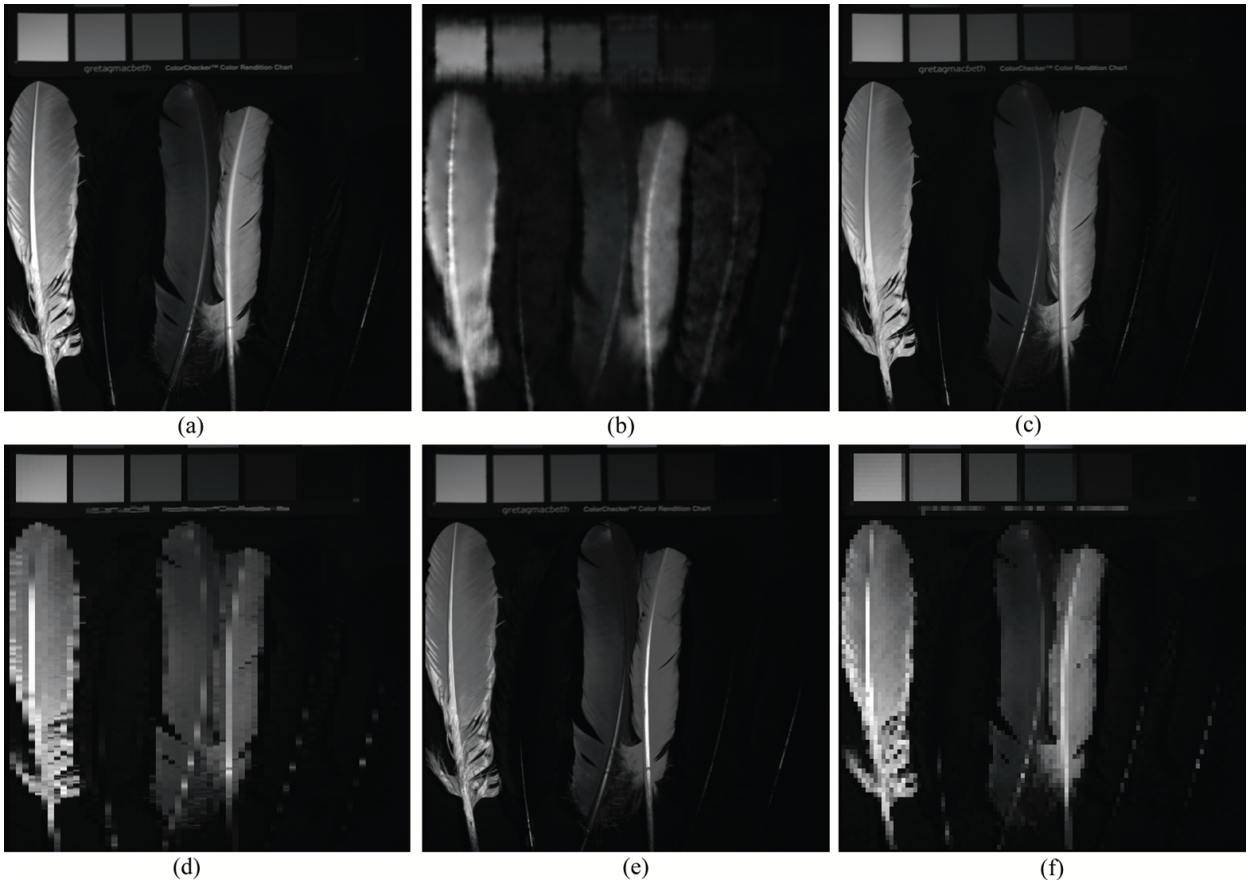


Fig. 11 Image comparisons at 560 nm which are captured by: (a) traditional spectrometer, (b) CASSI, (c) PMVIS, (d) LFIS, (e) training-based spectrometer, and (f) CQDs.

(4) LFIS and CQDs: The spatial resolution and spectral resolution in the light field imaging spectrometer and the CQDs presents a kind of tradeoff, and the increase of either leads to the decrease of the other. (5) The training-based spectrometers rely on strong assumptions and the accuracy of such is extremely sensitive to the training set.

To intuitively demonstrate the imaging quality, we utilized a public dataset (Yasuma *et al.*, 2010) to simulate the spectral images captured by different technologies. Noise and inference during 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 effectively recovered by CASSI. On the other hand, Fig. 11 shows the 2-D image at 560 nm, the spatial details in (b) are over-smoothed (which is captured by CASSI), and the spatial resolution of (d) and (f) is 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 the traditional remote sensing and the computer vision tasks (segmentation, material discrimination, photography, surveillance and so on), because of the remarkable efficiency in dynamic spectrum capture and simple low-cost system configurations. Meanwhile, several limitations also exist in different spectral imaging systems, which must account for further combinations of optical principles, CS theory, machine learning algorithms, and semiconductor device fabrication technologies.

References

Abed, Farhad Moghareh, Amirshahi, Seyed Hossein, Abed, Mohammad Reza Moghareh, 2009. Reconstruction of reflectance data using an interpolation technique. *JOSA A*, **26**(3):613–624.

Adelson, Edward H, Bergen, James R, 1991. The plenoptic function and the elements of early vision. MIT Press, p.3-20.

NASA, 2005. Deciphering ancient scrolls. Available from http://www.nasa.gov/vision/earth/technologies/scrolls.html

Arce, Gonzalo R, Brady, David J, Carin, Lawrence, Arguello, Henry, Kittle, David S, 2014. Compressive coded aperture spectral imaging: an introduction. *IEEE Signal Processing Magazine*, **31**(1):105–115.

Bao, Jie, Bawendi, Moungi G, 2015. A colloidal quantum dot spectrometer. *Nature*, **523**(7558):67–70.

Bioucas-Dias, José M, Figueiredo, Mário AT, 2007. A new twist: two-step iterative shrinkage/thresholding algorithms for image restoration. *IEEE Transactions on Image processing*, **16**(12):2992–3004.

Bodkin, Andrew, Sheinis, A, Norton, A, Daly, J, Beaven, S, Weinheimer, J, 2009. Snapshot hyperspectral imaging: the hyperpixel array camera. SPIE Defense, Security, and Sensing, p.73340H–73340H.

Boyd, Stephen, Parikh, Neal, Chu, Eric, Peleato, Borja, Eckstein, Jonathan, 2011. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning*, **3**(1):1–122.

Candès, Emmanuel J, Romberg, Justin, Tao, Terence, 2006. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, **52**(2):489–509.

Candès, Emmanuel J, Wakin, Michael B, 2008. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, **25**(2):21–30.

Cao, Xun, Du, Hao, Tong, Xin, Dai, Qionghai, Lin, Stephen, 2011a. A prism-mask system for multispectral video acquisition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **33**(12):2423–2435.

Cao, Xun, Tong, Xin, Dai, Qionghai, Lin, Stephen, 2011b. High resolution multispectral video capture with a hybrid camera system. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), p.297–304.

Cao, Xun, Yue, Tao, Lin, Xing, Lin, Stephen, Yuan, Xin, Dai, Qionghai, Carin, Lawrence, Brady, David J, 2016. Computational snapshot multispectral cameras. *IEEE Signal Processing Magazine*, **1053**(5888/16).

Chakrabarti, Ayan, Zickler, Todd, 2011. Statistics of real-world hyperspectral images. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), p.193–200.

Descour, Michael, Dereniak, Eustace, 1995. Computed-tomography imaging spectrometer: experimental calibration and reconstruction results. *Applied Optics*, **34**(22):4817–4826.

Descour, Michael, Volin, C.E., Ford, B.K., Dereniak, E.L., Maker, P.D., Wilson, D.W., 2001. Snapshot hyperspectral imaging. Integrated Computational Imaging Systems, p.IWB4.

Donoho, David L, 2006. Compressed sensing. *IEEE Transactions on Information Theory*, **52**(4):1289–1306.

Gao, Liang, Kester, Robert T, Hagen, Nathan, Tkaczyk, Tomasz S, 2010. Snapshot image mapping spectrometer (ims) with high sampling density for hyperspectral microscopy. *Optics Express*, **18**(14):14330–14344.

Gat, Nahum, 2000. Imaging spectroscopy using tunable filters: a review. AeroSense 2000, p.50–64.

Golbabaee, Mohammad, Vandergheynst, Pierre, 2012. Compressed sensing of simultaneous low-rank and joint-sparse matrices. *arXiv preprint arXiv:1211.5058*, .

Green, Robert O, Eastwood, Michael L, Sarture, Charles M, Chrien, Thomas G, Aronsson, Mikael, Chippendale, Bruce J, Faust, Jessica A, Pavri, Betina E, Chovit, Christopher J, Solis, Manuel, et al., 1998. Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (aviris). *Remote Sensing of Environment*, **65**(3):227–248.

Hao Du, Xin Tong, Xun Cao, Lin, Stephen, 2009. A prism-based system for multispectral video acquisition. 2009 IEEE 12th International Conference on Computer Vision, p.175-182. [doi:10.1109/ICCV.2009.5459162]

Harvey, Andrew R, Beale, John E, Greenaway, Alain H, Hanlon, Tracy J, Williams, John W, 2000. Technology options for imaging spectrometry. International Symposium on Optical Science and Technology, p.13–24.

Herrala, Esko, Okkonen, Jukka T, Hyvarinen, Timo S, Aikio, Mauri, Lammasniemi, Jorma, 1994. Imaging spectrometer for process industry applications. Optics for Productivity in Manufacturing, p.33–40.

Hunicz, Jacek, Piernikarski, Dariusz, 2001. Investigation of combustion in a gasoline engine using spectrophotometric methods. Optoelectronic and Electronic Sensors IV, p.307–314.

Kindzelskii, Andrei L, Yang, Zhi-yong, Nabel, Gary J, Todd, Robert F, Petty, Howard R, 2000. Ebola virus secretory glycoprotein (sgp) diminishes FcγRIIIB-to-CR3 proximity on neutrophils. *The Journal of Immunology*, **164**(2):953–958.

Kittle, David, Choi, Kerkil, Wagadarikar, Ashwin, Brady, David J, 2010. Multiframe image estimation for coded aperture snapshot spectral imagers. *Applied Optics*, **49**(36):6824–6833.

Lawlor, J, Fletcher-Holmes, DW, Harvey, AR, McNaught, AI, 2002. In vivo hyperspectral imaging of human retina and optic disc. *Investigative Ophthalmology & Visual Science*, **43**(13):4350–4350.

Liao, Xuejun, Li, Hui, Carin, Lawrence, 2014. Generalized alternating projection for weighted-2,1 minimization with applications to model-based compressive sensing. *SIAM Journal on Imaging Sciences*, **7**(2):797–823.

Lin, Xing, Liu, Yebin, Wu, Jiamin, Dai, Qionghai, 2014a. Spatial-spectral encoded compressive hyperspectral imaging. *ACM Transactions on Graphics (TOG)*, **33**(6):233.

Lin, Xing, Wetzstein, Gordon, Liu, Yebin, Dai, Qionghai, 2014b. Dual-coded compressive hyperspectral imaging. *Optics Letters*, **39**(7):2044–2047.

Ma, Chenguang, Cao, Xun, Wu, Rihui, Dai, Qionghai, 2014. Content-adaptive high-resolution hyperspectral video acquisition with a hybrid camera system. *Optics Letters*, **39**(4):937–940.

Mitchell, Peter A, 1995. Hyperspectral digital imagery collection experiment (hydice). Satellite Remote Sensing II, p.70–95.

Mooney, Jonathan M, Vickers, Virgil E, An, Myoung, Brodzik, Andrzej K, 1997. High-throughput hyperspectral infrared camera. *JOSA A*, **14**(11):2951–2961.

Morovic, Peter, Finlayson, Graham D, 2006. Metamer-set-based approach to estimating surface reflectance from camera rgb. *JOSA A*, **23**(8):1814–1822.

Morris, Hannah R, Hoyt, Clifford C, Treado, Patrick J, 1994. Imaging spectrometers for fluorescence and Raman microscopy: acousto-optic and liquid crystal tunable filters. *Applied Spectroscopy*, **48**(7):857–866.

Nguyen, Rang MH, Prasad, Dilip K, Brown, Michael S, 2014. Training-based spectral reconstruction from a single rgb image. European Conference on Computer Vision, p.186–201.

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

Rørslett, B, 2004. All you ever wanted to know about digital uv and ir photography. Available from http://www.naturfotograf.com/UV_IR_rev00.html

Schechner, Yoav Y., Nayar, Shree K., 2002. Generalized mosaicing: wide field of view multispectral imaging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**(10):1334–1348.

Shepp, Lawrence A, Vardi, Yehuda, 1982. Maximum likelihood reconstruction for emission tomography. *IEEE Transactions on Medical Imaging*, **1**(2):113–122.

Su, Lijuan, Zhou, Zhiliang, Yuan, Yan, Hu, Liang, Zhang, Siyuan, 2015. A snapshot light field imaging spectrometer. *Optik-International Journal for Light and Electron Optics*, **126**(9):877–881.

Wagadarikar, Ashwin A, Pitsianis, Nikos P, Sun, Xiaobai, Brady, David J, 2009. Video rate spectral imaging using a coded aperture snapshot spectral imager. *Optics Express*, **17**(8):6368–6388.

Willett, Rebecca M, Duarte, Marco F, Davenport, Mark A, Baraniuk, Richard G, 2014. Sparsity and structure in hyperspectral imaging: sensing, reconstruction, and target detection. *IEEE Signal Processing Magazine*, **31**(1):116–126.

Wu, Yuehao, Mirza, Iftekhar O, Arce, Gonzalo R, Prather, Dennis W, 2011. Development of a digital-micromirror-device-based multishot snapshot spectral imaging system. *Optics Letters*, **36**(14):2692–2694.

Wug Oh, Seoung, Brown, Michael S, Pollefeys, Marc, Joo Kim, Seon, 2016. Do it yourself hyperspectral imaging with everyday digital cameras. IEEE Conference on Computer Vision and Pattern Recognition, p.2461–2469.

Yamaguchi, Masahiro, Haneishi, Hideaki, Fukuda, Hiroyuki, Kishimoto, Junko, Kanazawa, Hiroshi, Tsuchida, Masaru, Iwama, Ryo, Ohyama, Nagaaki, 2006. High-fidelity video and still-image communication based on spectral information: Natural vision system and its applications. Electronic Imaging 2006, p.60620G–60620G.

Yasuma, Fumihito, Mitsunaga, Tomoo, Iso, Daisuke, Nayar, Shree K, 2010. Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum. *IEEE Transactions on Image Processing*, **19**(9):2241–2253.

Zhou, Zhiliang, Yuan, Yan, Xiangli, Bin, 2010. Light field imaging spectrometer: Conceptual design and simulated performance. Frontiers in Optics, p.FThM3.