nature portfolio

https://doi.org/10.1038/s41586-024-08109-1

Supplementary information

A broadband hyperspectral image sensor with high spatio-temporal resolution

In the format provided by the authors and unedited

Supplementary information: A broadband hyperspectral image sensor with high spatio-temporal resolution

Liheng Bian*†, Zhen Wang†, Yuzhe Zhang†, Lianjie Li, Yinuo Zhang, Chen Yang, Wen Fang, Jiajun Zhao, Chunli Zhu, Qinghao Meng, Xuan Peng and Jun Zhang*

State Key Laboratory of CNS/ATM & MIIT Key Laboratory of Complex-field Intelligent Sensing, Beijing Institute of Technology, Beijing 100081, China.

*Corresponding author(s). E-mail(s): bian@bit.edu.cn; zhjun@bit.edu.cn;
†These authors contributed equally to this work.

Contents

1 Comparison of different snapshot hyperspectral imaging techniques

We briefly review the existing snapshot hyperspectral imaging techniques and compare them with our work in terms of spatial resolution, spectral resolution, and spectral range to clarify the differences in the Extended Data Tab. 1. Based on different integration types, we classify the snapshot spectral imaging approaches into two categories: (i) system integration and (ii) on-chip integration. System integration employs separate optical components for hyperspectral imaging. Although these systems effectively improve temporal resolution, achieving broadband high-resolution hyperspectral imaging remains challenging, resulting in poor universality and flexibility. On-chip integration can offer more compact and potentially lower-cost solutions. However, existing on-chip techniques suffer from low resolution and low light throughput, which prevents their practical applications in high-resolution hyperspectral imaging.

The main contribution of this work lies in reporting a low-cost on-chip computational hyperspectral imaging architecture with high light throughput and high resolution over a wide wavelength range. To validate the effectiveness of this architecture, we fabricated two visible-near-infrared (VIS-NIR) range hyperspectral sensors: HyperspecI-V1 (400-1000 nm) and HyperspecI-V2 (400-1700 nm). The acquisition frame rate of HyperspecI-V1 is 47 frames per second (fps), with a reconstruction frame rate of 31 fps (61 channels with 2048×2048 pixels). For HyperspecI-V2, the acquisition fps is 134, with a reconstruction frame rate of 105 fps (96 channels with 1024×1024 pixels). These sensors are capable of real-time acquisition in the VIS-NIR spectral range, making them a readily available alternative to existing cameras for various applications.

2 Broadband spectral modulation materials

Under the proposed HyperspecI architecture, we have initially investigated the possible available material systems for BMSFA, including semiconductor materials (nano-metal oxides, quantum dots), organic dyes, carbon-based nanomaterials, nano noble metals (gold nanorods), etc. However, by a preliminary demonstration of the experiment, we found that some of the materials are not suitable or cannot produce the optimal performance. Specifically, through the literature study and our experiment, we observed challenges in achieving the integration of quantum dots into micro-nano arrays and fluorescence quenching, resulting in suboptimal modulation effects. Moreover, carbon-based nanomaterials face challenges in realizing arrays with diverse light modulation characters, and integration proves challenging, with limitations in achieving a broad working wavelength. Similarly, nano noble metals exhibit aggregation during preparation and fabrication, posing an impact on light modulation performance. We further compared the specific performance of several potential material systems in broadband spectral modulation through material preparation and testing. Obtain a large number of materials that can be used for broadband spectral modulation as basic candidate materials through experiments and large-scale preparation. Then we selected the optimal combination from the candidate materials through the proposed material-optimized selection method for preparing BMSFA. The specific experimental process is shown below.

2.1 Material preparation

Through a comprehensive exploration of the existing literature and our accumulated research experience in light modulation materials, we have gained valuable insights into the spectral modulation effects within the visible and near-infrared bands. We learned that carbon nanomaterials [S1, 2], quantum dot materials [S3], organic dyes, nano-metal oxides, nano-precious metals [S4] exhibit potential effective spectral modulation effects in the visible and near-infrared band. Considering the spectral modulation properties, material size, stability, low cost large-scale preparation capability and CMOS process compatible of the aforementioned materials, we focused on GNDs, nano-metal oxides, and organic dyes. Nano-precious metals emerge as promising materials owing to their outstanding light modulation effects. Specifically, gold nanospheres and gold nanorods (GNDs) exhibit controllable transmission spectrum ranging from 500 nm to 1800 nm, making them possiable to meet our requirements. Nano-metal oxides and composites doped with rare-earth metals exhibit a remarkable light modulation effect despite the challenge of controllable modulation, such as tungsten oxide [S5], tin antimony oxide [S6], cesium tungsten bronze [S7], etc. Moreover, their stability and particle size also meet our expectations. Organic dyes stand out as visible and near-infrared light modulation materials due to their stable chemical properties and significant absorption signature, such as cyanine dyes [S8], phthalocyanine dyes [S9], quinone dyes [S10], azo dyes [S11], metal complex dyes [S12], etc. These organic materials exhibit a good controllability of spectral modulation and well-established production process. Additionally, their solubility in common organic solvents facilitates integration into semiconductor processing workflows.

The following are experiments conducted for these three materials.

1. GNDs

GNDs are rod-shaped nanoparticles composed of gold, characterized by adjustable absorption peaks and stable chemical properties. Their absorption peak can be shifted from 500 nm to 1800 nm by adjusting their aspect ratio [S13]. In this study, we fabricated oil-soluble GNDs with an absorption peak of 1600 nm and a concentration of 0.1 mg/ml. However, the spectral modulation effect of GNDs was not observed when the spin-coated material reached several micrometers, as their transmission remained consistently above 90%.

To further evaluate the feasibility of GNDs, we utilized centrifugation to concentrate the GNDs solution, leading to a five-fold increase in concentration. However, the transmission spectra of the concentrated GNDs solution remained around 90%, comparable to the previous concentration. Despite our efforts to further increase the concentration, we faced sedimentation issues at higher concentrations.

2. Organic dyes

We validated the spectral properties of 23 organic dyes using spectrophotometer, after curing them on a quartz substrate with a spin coating thickness of 10 μm. The 23 types of organic dye (OD) are named as OD-480, OD-520, OD-560, OD-580, OD-590, OD-600, OD-610, OD-620, OD-630, OD-660, OD-680, OD-700, OD-740, OD-780, OD-800, OD-810, OD-820, OD-830, OD-840, OD-890, OD-920, OD-950, OD-970. These organic materials exhibit remarkable light modulation properties at specific wavelengths, as shown in Extended Data Fig. 9d, mainly comprising cyanine dynes, phthalocyanine dyes, quinone dyes, azo dyes, metal complex dyes, etc.

In order to prepare them into photolithography compatible materials, we directly mixed 0.2 g of each organic dye with 20 ml of photoresist and utilized an ultrasonic liquid processor (NingHuai NH-1000D) at room temperature. To ensure complete dissolution and remove impurities, the mixed solution was filtered using 3 μm pore size filters. Subsequently, we applied the mixed solution onto the quartz substrates, and the test smears formed at 4000 rpm on the spin coater. Then, we measured the transmission spectra of these smears using spectrophotometer (PerkinElmer Lambda 950) at 200-1700 nm. As shown in Extended Data Fig. 9d, the organic dyes exhibit prominent light modulation spectra between 400-1200 nm. This confirms the exceptional suitability of organic dyes for achieving spectral modulation within the 400-1200 nm range.

3. Nano-metal Oxides

We prepared and tested the broadband spectral modulation characteristics and preparation compatibility of nano-metal oxides. These nano-metal oxides include blue tungsten oxide ($WO_{2.90}$) [S5], cesium tungsten bronze (Cs_xWO_3) [S7], antimony tin oxide [S6] (ATO), samarium oxide (Sm_2O_3) [S14], cerium oxide (CeO_2) [S15] and yttrium oxide (Y_2O_3) [S16]. We individually mixed 0.5 g of each material powder with 3 ml of photoresist and stirred them using an ultrasonic liquid processor for 15 minutes at room temperature. Subsequently, we evaluated the uniformity of each mixture to assess their applicability. The unfiltered solution displayed the color of the material powder, while the filtered solution appeared significantly lighter. This could be attributed to the material powder aggregating in the photoresist, forming larger particles that were subsequently filtered out. The filtered samples failed to produce any spectral modulation effect, as the spectral curves closely resembled those of pure photoresist smears, indicating a low concentration of filtered solutions.

To further evaluate the modulation capability of the nano-metal oxides, we decided to increase the concentration and avoid filtering the solutions. We individually mixed 1 g of each material powder with 3 ml of the photoresist, and stirred the mixture using the ultrasonic liquid processor at room temperature for 15 minutes. the resulting smears exhibited a substantial thickness. The transmission of the smears in the near-infrared band was nearly negligible, while the transmission of the photoresist exceeded 90%, indicating an excessively high concentration.

We optimized the thickness of the smears. Furthermore, we employed diffuse reflection measurement to address the light blockage issue caused by the increased thickness of the smears. Although $WO_{2.90}$, Cs_xWO_3 and ATO materials exhibit some light modulation characters in the near-infrared band, the spectral curves lack significant modulation signature. On the other hand, Y_2O_3, CeO_2, Sm_2O_3 demonstrate spectral curves remarkably similar to that of the sample holder, indicating minimal light absorption. The dark-colored materials, such as $WO_{2.90}$, Cs_xWO_3, ATO, exhibit pronounced light absorption compared to the light-colored materials, such as Y_2O_3, CeO_2, Sm_2O_3. Hence, nano-metal oxides can provide light spectral modulation effects in the near-infrared band, with dark-colored materials possessing higher absorption capacities than the light-colored materials.

In order to disperse nano-metal oxides at appropriate concentrations, we introduced dispersants. These dispersants effectively scatter the nano-metal oxide powder, reducing the formation of clumped particles and promoting the creation of a stable material dispersion fluid. This approach addresses the issues of filtration and concentration by diluting the solution. A widely used dispersant in the industry is Propylene Glycol Methyl Ether Acetate (PGMEA), which functions as both a dispersion agent and a diluent for photoresist. PGMEA facilitates the dispersion of the nano-metal oxide powder while simultaneously aiding its integration with the photoresist.

We mixed 40 g of each material powder with 160 g of PGMEA respectively. The mixture underwent dispersion for 48 hours using the ultrasonic liquid processor at room temperature, resulting in material dispersion fluids with a mass fraction of 20%. To determine the optimal concentration, we prepared 5 concentration gradients for ATO, Sm_2O_3, and CeO_2. For each material, we mixed 3 ml of the corresponding material dispersion fluid with 3 ml, 6 ml, and 9 ml of photoresist. The mixture was then stirred for 15 minutes using the ultrasonic liquid processor at room temperature, yielding material solutions with volume ratios of 1:1, 1:2, and 1:3. Additionally, we mixed 3 ml of each material dispersion fluid with 0.2 g and 0.4 g of material powder, along with 3 ml of photoresist. After stirring the mixture for 15 minutes using the ultrasonic liquid processor, we obtained each material solution with volume ratios of 1:1 and respective material powder amounts of 0.2 g and 0.4 g. Spectrophotometers were employed to measure the samples of the aforementioned five concentrations (Extended Data Fig. 9b). The samples and their corresponding results are presented separately in Extended Data Fig. 9e-f. It was observed that CeO_2, Sm_2O_3 did not exhibit spectral modulation capability within the range of 1000 nm to 1700 nm under a coating thickness of approximately a dozen microns. ATO showed varied spectral curves within the 1000-1700 nm range, indicating a favorable spectral modulation capability at concentrations of 1:1 and 1:2. This experiment verifies that light-colored materials possess limited spectral modulation capability under a coating thickness of a dozen microns, whereas dark materials perform significantly better. Therefore, light-colored materials do not meet our requirements.

We further conducted tests on dark metal oxides in close proximity to Indium, Tin, and Tungsten. These nano-metal oxides include CoO, Pr_6O , TiO_2/Ag, CuO, ZnO, Fe_2O_3, NiO, Nd_2O_3, MnO and Cr_2O_3. We mixed 40 g of each material powder with 160 g PGMEA respectively. Following a dispersion process of 48 hours using the ultrasonic liquid processor at room temperature, we obtained material dispersion fluids with a mass fraction of 20%. Subsequently, we mixed 3 ml of dispersion fluids from 12 different materials (including the newly introduced materials along with $WO_{2.90}$, Cs_xWO_3) with 3 ml photoresist, respectively. The mixture was stirred for 15 minutes using the ultrasonic liquid processor at room temperature. The variation of MnO and Cr_2O_3 in the curves exhibited satisfying modulation capability. However, the concentration was deemed unsuitable for 7 materials, namely CoO, TiO_2/Ag, ZnO, NiO, Nd_2O_3, Cs_xWO_3 and PrO_3. Therefore, multiple concentrations need to be configured. The remaining materials yielded unfavorable results. To address the issue of inappropriate concentration, we mixed 3 ml of each material dispersant with 3 ml, 6 ml, and 9 ml photoresist separately. These mixtures were stirred for 15 minutes using the Ultrasonic liquid processor at room temperature. The outcomes are displayed in Extended Data Fig. 9e-f. Notable modulation effects were observed for TiO_2/Ag, Cs_xWO_3, CoO at a concentration ratio of 1:2.

2.2 Optimized material selection

We have an evolutionary optimization method to determine the filter number and select the optimal materials for BMSFA fabrication, as detailed below.

1. The evolutionary optimization method for filter selection and BMSFA design.

The design of BMSFA for light modulation is crucial for high-precision spectral reconstruction. The key factors in BMSFA design include the transmission spectra of materials and the number of modulation filters. The BMSFA design process includes the following steps: initial coarse selection of spectral modulation material system, fine-grained selection of spectral modulation materials, and numerical simulation of BMSFA.

Within the material system of organic dyes, thousands of materials with different transmission spectra in the spectral range of 400-1200 nm can be obtained through molecular structure design [S17–19]. For nano-metal oxides, hundreds of materials with different transmission spectra in the near-infrared range can be obtained by changing their particle size, shape, and dispersion concentration [S20]. Considering low large-scale manufacturing cost and high light modulation performance, we prepared ∼50 organic dyes and ∼20 nano-metal oxides. With further considerations after curing the materials on a quartz substrate with a spin coating thickness of 10 µm in our BMSFA fabrication, we finally determined candidate 23 organic dyes and 12 nano-metal oxides with appropriate spectral characteristics as shown in Extended Data Fig. 8d. Next, we developed the evolutionary optimization method to select materials for experiment fabrication of BMSFA.

Second, we focus on the fine-grained selection of spectral modulation materials among the two material systems, which involves selecting the optimal filter combination from the 23 organic dyes and 12 nano-metal oxides mentioned above to form the filter unit of BMSFA. We developed an evolutionary optimization based method [S21] to choose specific materials to achieve the highest reconstruction precision. The schematic of this method is presented in Extended Data Fig. 8a, which includes the following steps. The computational complexity of such an evolutionary optimization based material selection method is $O(n)$.

(1) Construct a large-scale hyperspectral image dataset, which includes spectral data of various scenes and illumination conditions in the spectral range of 400-1700 nm.
(2) Compress the spectral image dataset along the spectral dimension using the PCA technique [S22] to obtain data distribution maps and reconstruction errors (as shown in Extended Data Fig. 8e) in different compressed low dimensional spaces. Analyze the spectral sparsity of hyperspectral data and define a potential low dimensional spectral space with dimension n (because observation spaces of different dimensions have different reconstruction accuracy, we define dimensions within a range that constitutes a set of dimensions).
(3) Sequentially select different spectral dimensions as the number of modulation materials and select the materials based on this number.
(4) Start material selection based on evolutionary optimization. Select a subset of spectral modulation materials from the material database as the initial selection, and then replace some of the materials through survival of the fit, crossover, mutation, and random replacement operations to update a new subset of materials. Continuously repeat the above process until the reconstruction accuracy using the updated material subset converges to the best level in the aforementioned spectral dataset. In our implementation, we used the CVX-ECOS solver for spectral reconstruction, and employed spectral fidelity to evaluate reconstruction accuracy.
(5) Repeat (3) - (4) steps to evaluate the spectral reconstruction accuracy for different filter numbers, and plot the relationship map between spectral reconstruction accuracy and the subset of materials to determine the optimal number of filters.

2. Explanation of the rationality of the determined number and selection of filters.

Through the above material selection process, we obtained the optimal filter combinations of different filter numbers. We further evaluated the spectral reconstruction accuracy under different number of materials (i.e. filters).

As shown in Extended Data Fig. 8d, we can see that in the range of 400-1000 nm, the spectral accuracy tends to converge when there are more than 5 organic dye filters, and there is almost no significant change of spectral accuracy when the number is greater than 12. For the range of 1000-1700 nm, the spectral accuracy gradually converges when there are more than 4 nano-metal oxide filters, and the accuracy remains almost unchanged when the number is greater than 7. Therefore, considering the tradeoff between spectral accuracy and spatial resolution of the BMSFA, we selected 10 organic dyes and 6 nano-metal oxides to achieve an approximately optimal selection. Together with the powerful SRNet algorithm, such optimized BMSFA achieved high-resolution hyperspectral imaging in the wide spectral range of 400-1700 nm.

3. The performance of the BMSFA using the selected filters.

Using the above selected filters, we evaluated the spectral accuracy of this BMSFA under different measurement noise levels. Table S1 shows the spectral accuracy under different input SNRs with the selected 10 organic dyes and 6 nano-metal oxides at different spectral ranges. We can see that the BMSFA produced high spectral fidelity even in low SNR conditions.

Tab S1: The spectral accuracy under different input SNRs with the selected 10 organic dyes and 6 nano-metal oxides at different spectral ranges.

Input SNR	Spectral range		
	400-1000nm	1000-1700nm	400-1700nm
INF	99.20%	98.82%	98.97%
30 dB	98.32%	97.93%	98.21%
20 dB	97.97%	97.09%	97.53%
10 dB	93.53%	92.37%	92.64%

The spectral modulation materials to fabricate the BMSFA in our prototype are composed of organic dyes and nano-metal oxide materials. Due to the efficient manufacture and preparation of organic dyes, we achieved effective and controllable spectral modulation in the spectral range of 400-1200 nm. As the manufacture of organic dyes above 1200 nm is challenging, we further prepared nano-metal oxide materials that played a dominant role for spectral modulation at 1200-1700 nm. Besides, following the above method, we selected the optimized filter number and the corresponding materials from the two material systems.

We analyzed the effective dimensions of different spectral ranges using the data dimensionality reduction technique, as shown in Extended Data Fig.8 and Tab.S1. We measured 100,000 sets of realistic hyperspectral data in the range of 400-1000 nm and 1000-1700 nm. For spectra with a sampling rate of 10 nm in the range of 1000-1700 nm, we can achieve a spectral accuracy of over 95% using 5 spectral modulation materials. In the spectral range of 400-1000 nm, at least 6 spectral modulation materials are required to achieve this accuracy. Our BMSFA design coincides with the analysis, leading to satisfactory hyperspectral imaging performance for both visualization and various applications.

3 HyperspecI fabrication

3.1 BMSFA fabrication

This section focuses on the fabrication of the broadband multispectral filter array (BMSFA), which includes a series of processes of photoresist dissolution, photomask design, substrate preparation, photoresist coating, soft bake, UV exposure, post-exposure bake, development, and hard bake. The BMSFA is a crucial component of our HyperspecI sensor, which modulates the spectra of incident light. Firstly, materials with different spectral modulation properties were dissolved in photoresists after filtration. This resulted in photoresists with different spectral modulation characteristics. Subsequently, we dropped a solvent of one kind of photoresist onto the quartz substrate, ensuring uniform distribution of the photoresist containing spectral modulation materials using a spin coater (Helicoater, HC220PE). Next, we employed UV photolithography to cure the photoresist containing a kind of modulation material at the special position of the quartz substrate. This process was repeated for different modulation materials using a designed photomask. Finally, we poured pure SU-8 photoresist onto the finished substrate, completing the photolithography process through photoresist coating, soft bake, UV exposure, post-exposure bake, and hard baking. The specific steps are as follows:

Step 1: We mixed 16 spectral modulation materials with photoresist at a rate of 10 mg/ml for each material. The mixed photoresist solution was allowed to stand and then filtered using 3 µm pore size filters to remove undissolved impurities. This resulted in 16 types of photoresists with distinct spectral modulation characteristics.

Step 2: The quartz substrate was secured onto the spin coater, and the photoresist was evenly dropped onto the substrate using a pipette in a spiral outward manner from the center for diffusion. After applying an adequate amount of photoresist, the spin coater was set to a uniform rise speed of 10 s up to 4000 rad/s, followed by a 60 s uniform rotation for leveling. This process ensured uniform attachment of the material to the substrate with consistent thickness. After photoresist coating, we put the substrate on a hot plate at 95°C for a soft bake of 5 minutes to enhance the adhesion of the photoresist. Following the soft bake, the substrate was positioned in the photolithography machine (SUSS MA6 Mask Aligner, SUSS MicroTec AG). We then installed the designed photomask and selected the appropriate microscope head. As shown in Extended Data Fig. 7d, the photomask consists of a series of basic structural units with a size of 10 µm, and the distance between different units is 40 µm. The exposure dose of the UV lithography machine is 1,000 mJ/cm^2. After UV exposure, we placed the substrate back on the 95°C hot plate for a post-exposure bake of 10 minutes. The step in negative photoresist (SU-8) lithography aimed to enhance crosslinking in the exposed areas and promote the removal of unexposed regions. Additionally, post-exposure bake can also reduce the standing wave effect and improve overall resolution and fidelity of the desired pattern. Then, we placed the substrate in the developer to remove the photoresist from the unexposed portion. After development, the developer solution is quickly dried using compressed air. Finally, the substrate was placed on a hot plate for a hard bake at 150°C for 5 minutes, followed by natural cooling. This resulted in a substrate engraved with the desired material in the designed positions.

Step 3: The substrate was placed back in the spin coater for the application of the second photoresist containing a different spectral modulation material. The photoresist coating and soft bake were repeated. Subsequently, the substrate was aligned with the slot mark to achieve pre-alignment of the lithography machine. The microscope was then utilized to locate and focus the alignment mark. Afterward, we adjusted the X/Y direction or underwent θ angle rotation to align the photomask version with the substrate. The mask alignment table was adjusted to position the photomask precisely, and the vacuum exposure mode was selected. Following exposure, the process of post-exposure bake, development, and hard bake as in step 2, was repeated to cure the spectral modulation materials onto the substrate. This process was replicated to complete the photolithography process for the remaining 14 materials.

Step 4: After completing the photolithography process with 16 different spectral modulation materials on the substrate, we filled the gaps between different material points using pure SU-8 photoresist to ensure a smooth surface. We utilized the flood exposure mode of the SUSS MA6 photolithography machine to expose the substrate with SU-8 photoresist for 60 s, followed by the hard bake process to solidify the surface. This step effectively prevents the concentration of light caused by the unevenness of the lithographic substrate.

By following the aforementioned steps, the BMSFA was prepared for our hyperspectral imaging sensor. We constructed a 4 × 4 array arrangement of the BMSFA with each material unit of 10 µm in size, incorporating 16 distinct broadband spectral modulation materials spanning the VIS-NIR spectral range. For BMSFA production, we utilized 4-inch substrates, each containing 40 low-cost BMSFAs (Extended Data Fig. 7e).

3.2 HyperspecI sensor integration

The spectral modulation effect of our sensor is achieved by integrating a BMSFA onto the sensor surface. Placing the BMSFA on the image plane and relaying it to the sensor plane using a bulky relay lens is a commonly used but cumbersome method. It often leads to severe vignetting in large field-of-view imaging. In order to achieve the best integration of our HyperspecI sensor, we adopted a different approach. We utilized a laser engraving machine to remove the packaging glass from the monochrome sensor. Then we cured the BMSFA onto the sensor surface using photoresist under UV lighting, ensuring optimal sensor integration. Extended Data Fig. 7b-c shows the overall information of the integrated sensor.

4 Modeling and calibration of HyperspecI

4.1 The modeling of HyperspecI

HyperspecI sensor follows the principle of snapshot compressive sensing imaging [S23, 24], where the spatial-spectral data cube from the target scene is encoded through the BMSFA and quantized for acquisition as 2D measurement on the monochrome image sensor. This process can be described as a linear process of degradation of 3D data into 2D, where BMSFA acts as a linear degradation operator and also as a measurement matrix in compressed sensing theory. The compressed sensing matrix represents the essential characteristics of the sensor on spatial-spectral modulation, which is the key to the sampling process modeling and hyperspectral image reconstruction. The forward process is expressed in mathematical form as follows:

$$b(h,w) = \int_{\lambda_1}^{\lambda_2} r(h,w,\lambda) d(h,w,\lambda) c(h,w,\lambda) d\lambda \tag{S1}$$

where $b(h,w)$ denotes the measurement captured by the sensor at the (h,w) position pixel, $d(h,w,\lambda)$ denotes the spectral transmissivity of BMSFA at position (h,w) and wavelength λ, and $c\,(h,w,\lambda)$ denotes spectral response of optical systems such as monochromatic sensors and camera lenses. By discretizing the above process, we can get

$$\mathbf{B}_{h,w} = \sum_{k=1}^{M} \mathbf{R}_{h,w,k} \odot \mathbf{D}_{h,w,k} \odot \mathbf{C}_{h,w,k} + \mathbf{N}_{h,w} \tag{S2}$$

where $\odot$ denotes element-wise multiplication, $\mathbf{N}_{h,w}$ denotes the measurement noise, and k is the discrete spectral wavelength, M is the number of the spectral channels. After vectorizing $\mathbf{b} = \mathrm{vec}(\mathbf{B})$, $\mathbf{r} = \mathrm{vec}(\mathbf{R})$, $\mathbf{H} = \mathrm{vec}(\mathbf{D} \odot \mathbf{C})$, and $\mathbf{n} = \mathrm{vec}(\mathbf{N})$, Eq. (S2) can be expressed as

$$\mathbf{b} = \mathbf{H}\mathbf{r} + \mathbf{n} \tag{S3}$$

where $\mathbf{H}$ denotes the sensing matrix of the BMSFA camera. The accurate measurement of the sensing matrix $\mathbf{H}$ is a necessary condition for achieving high-precision spectral reconstruction. Although the spectral responses of BMSFA, monochrome image sensors, and imaging lenses were considered separately during the design stage, the $\mathbf{H}$ matrix synthesized solely through numerical simulation has a large error, and thus requires actual calibration. Considering that calibration is a tedious and time-consuming process, and in order to achieve automated calibration of multiple sensors, we designed an automatic calibration system based on Omno151 monochromator that can automatically obtain the compressed sensing matrix $\mathbf{H}$ of the sensors to be calibrated by only placing them in the calibration light path.

4.2 The calibration of HyperspecI

As shown in Extended Data Fig. 7f-g, we used a monochromator (grating monochromator, Omno151, working range 200-2000 nm) to generate monochromatic light with a full width at half maximum (FWHM) of 10 nm at the target spectral wavelength. The monochromatic light was uniformly irradiated onto a power meter probe (Thorlabs S130VC, S132C) and the surface of the HyperspecI sensor after passing through a collimated optical path. In the automated calibration process, a program we developed was used to control the output wavelength of the monochromatic light from the monochromator, acquire the power meter value, and save the corresponding two-dimensional response values of HyperspecI sensor under the certain wavelength of monochromatic light. This process was repeated for each wavelength from λ_1 to λ_2 to eventually automatically generate the measurement matrix $\mathbf{H}$.

4.3 The performance of sensor chip

To be more specific, we employed the SONY IMX264 and SONY IMX990 chips to independently achieve hyperspectral imaging in the spectral ranges of 400-1000 nm (HyperspecI-V1) and 400-1700 nm (HyperspecI-V2). The performance analyses were conducted for each sensor as follows.

(1) The IMX264 chip is silicon-based with a spectral range of 400-1000 nm. It has a spatial resolution of 2448 $\times$ 2048 pixels (we used the central 2048 $\times$ 2048 pixels for hyperspectral imaging), a maximum frame rate of 47 frames per second (8-bit), utilizes a global shutter, features a 2/3-inch sensor size, and has a pixel size of 3.45 µm $\times$ 3.45 µm. The dynamic range is 72 dB.

Using the IMX264 chip, our HyperspecI-V1 sensor obtains 61 spectral channels with a 10 nm interval over the 400-1000 nm range. Utilizing HypersepcI-V1, we conducted hyperspectral reconstructions in both outdoor and laboratory settings, as illustrated in Fig. 2a of the Main Manuscript and Extended Data Fig. 3. To further validate the sensor's performance, we conducted multiple application experiments including fruit sugar detection (with characteristic peaks at 680 nm and 970 nm [S25–27]), chlorophyll content detection (with characteristic peaks at 660 nm and 720 nm [S28]), and blood oxygen detection (with characteristic peaks at 780 nm and 830 nm [S29], as shown in Supplementary Section 7 and 8.

(2) The Sony Indium Gallium Arsenide chip IMX990 further extends the spectral range to 400-1700 nm. This chip has a 1/2-inch sensor size, a pixel size of 5 µm $\times$ 5 µm, a spatial resolution of 1280 $\times$ 1024 pixels (we used the central 1024 $\times$ 1024 pixels for hyperspectral imaging), and a dynamic range of 51.5 dB.

Using the Sony IMX990 chip, we fabricated the HyperspecI-V2 sensor, enabling hyperspectral imaging spanning from 400 to 1700 nm. To the best of our knowledge, it currently stands as the most expansive spectral detection range achievable through a single detector within the VIS-NIR range. Even mechanical scanning hyperspectral cameras based on grating or prism are limited to achieve such wide spectral detection through a single detector (typically requiring the stitching of different sensors and gratings) [S30–32]. Utilizing the HypersepcI-V2 sensor, we conducted spectral reconstructions in both outdoor and laboratory settings, as depicted in Fig. 2a of the Main Manuscript and Extended Data Fig. 3. To further validate the performance of this sensor, we conducted textile fabric identification and apple bruise detection experiments, as illustrated in Fig. 6 of Main Manuscript. The cotton fabrics exhibit prominent feature peaks/valleys around 1220, 1320, and 1480 nm, while the polyester fabrics demonstrate distinct feature peaks/valleys near 1320, 1420, and 1600 nm [S33, 34]. Apples exhibit distinctive spectral characteristics near wavelengths of 1060 nm, 1260 nm, and 1440 nm, which are particularly crucial for invisible bruising detection. The experiment details are referred to Supplementary Section 9.

5 Deep-learning based reconstruction

5.1 Dataset construction

Our dataset was mainly captured using the commercial FigSpec-23 and GaiaField Pro-N17E-HR hyperspectral cameras, both integrated under a push-broom scanning mechanism, as shown in Extended Data Fig. 5a. The Figspec-23 camera covers a spectral range of 400-1000 nm, with a spectral resolution of 2.5 nm, and captures HSIs at a spatial resolution of 960 × 1230 pixels. After data processing, 61 bands were selected for our dataset at 10 nm intervals in the 400-1000 nm range. The GaiaField Pro-N17E-HR camera spans a spectral range of 900-1700 nm, with a spectral resolution of 5 nm, capturing images at a spatial resolution of 640 × 666 pixels. After data processing, 81 bands were selected for our dataset at 10 nm intervals in the 900-1700 nm range.

Because the field of view of these two cameras is different, We performed image registration on the acquired hyperspectral images of the two cameras, using the scale-invariant feature transform (SIFT) technique [S35–37]. As shown in Extended Data Fig. 5b, this process finds the similar spatial features of the two cameras, using which to register the field of view. As a result, there yields the hyperspectral imaging dataset covering the entire spectral range of 400-1700 nm, with a spatial resolution of 640 × 666 pixels and a total number of 131 spectral bands at 10 nm intervals.

As shown in Extended Data Fig. 5c, an exemplar visualization of the constructed hyperspectral imaging dataset is presented. We captured 1000 sets of HSIs using the Figspec-23 and GaiaField Pro-N17E-HR hyperspectral cameras. This dataset includes 500 outdoor scenes (such as road, building, flower, vehicle, etc.) and 500 indoor scenes under laboratory lighting (including food, cloth, toy, model, etc.). The outdoor scenes were photographed in Beijing, China from April to June and October to December, encompassing various weather conditions including sunny, cloudy, and overcast.

During our SRNet model training, considering the gradual spectral variations in the 1000-1700 nm range, we reduced the number of spectral channels in this range to improve both training and testing efficiency. Consequently, the training data comprised 96 spectral channels, with 61 channels at intervals of 10 nm in the 400-1000 nm range and 35 channels at intervals of 20 nm within the 1000-1700 nm range.

The raw data acquired using scanning hyperspectral imaging systems (FigSpec-23 and GaiaField Pro-N17E-HR) contain the Quantum Efficiency (QE) of the sensor, introducing bias compared to the real spectra. This condition is consistent with prevailing commercial hyperspectral cameras such as Specim [S38], Headwall [S39], Figspec [S40], and spectrometers like Thorlabs [S41] and Ocean Optics [S42]. Notably, these instruments directly capture spectral data in Digital Numbers (DN), influenced by the QE of detectors and optical systems.

Tab S2: Overview of the existing hyperspectral imaging datasets.

Ref.	**Name**	**Scene number**	**Spectral range**	**Channel number**	**Spatial resolution**	**Data Type**
[**S43**]	ICVL	200	400-700 nm	31	1392×1300	Raw Data
[**S44**]	Harvard	50	420-720 nm	31	1392×1024	Raw Data
[**S45**]	CAVE	32	400-700 nm	31	512×512	Reflectance Data
[**S46**]	KAIST	30	400-700 nm	31	3376×2704	Reflectance Data
[**S47**]	Manchester	30	400-720 nm	33	1344×1024	Reflectance Data
[**S48**]	TokyoTech 59	16	420-1000 nm	59	512×512	Reflectance Data
[**S49**]	TokyoTech 31	30	420-720 nm	31	2048×2048	Reflectance Data
[**S50**]	ARAD	1000	400-700 nm	31	512×482	Radiometric Data
Ours	**HyperspecI-V1**	**1000**	**400-1000 nm**	**61**	**1230×960**	**Radiometric Data**
Ours	**HyperspecI-V2**	**1000**	**400-1700 nm**	**131**	**666×640**	**Radiometric Data**

Also, we would like to note that some public HSI datasets also contain the hyperspectral camera's QE, such as ICVL [S43], Harvard [S44], etc. For computational hyperspectral imaging systems, the reconstructed HSI constituents are greatly influenced by the HSI dataset used for network training [S57, 58].

Tab S3: The training datasets used in different computational hyperspectral imaging systems.

Ref.	Dataset name	Scene amount	Spectral range	Channel number	Spatial resolution of training data
[S51]	ICVL, CAVE	232	400 - 700 nm	31	1392×1300, 640×480
[S52]	CAVE, KAIST	62	450 - 650 nm	28	512×512, 3376×2704
[S53]	CAVE, KAIST	62	450 - 650 nm	28	512×512, 3376×2704
[S54]	CAVE, KAIST	62	406 - 684 nm	12	512×512, 3376×2704
[S55]	Harvard, ICVL, KAIST	238	420 - 660 nm	25	1392×1024, 1392×1300, 3376×2704
[S56]	CAVE, Manchester, TokyoTech 31	92	450 - 650 nm	20	512×512, 1344×1024, 2048×2048
Ours	**HyperspecI-V1**	**1000**	**400 - 1000 nm**	**61**	**1230×960**
Ours	**HyperspecI-V2**	**1000**	**400 - 1700 nm**	**131**	**666×640**

As such, we summarized the existing public hyperspectral datasets in Tab. S2. The available datasets are categorized into three types, including raw data, reflectance data, and radiometric data.

- **Raw Data.** This kind of dataset (such as ICVL [S43] and Harvard [S44]) contains the raw digital numbers (DN) data acquired by hyperspectral cameras, corresponding to the multiplication of the illumination spectrum and the camera's QE. As such, these datasets contain both illumination and QE [S59, 60].
- **Reflectance Data.** This kind of dataset (such as CAVE [S45], TokyoTech 31 [S49], TokyoTech 59 [S48], and KAIST [S46]) removes the influence of illumination and QE by reflectance calibration. The common strategy utilizes a whiteboard (typically composed of barium sulfate) and the hyperspectral imaging system to obtain the baseline reference. Subsequently, the target spectra undergo a division operation with this baseline reference to produce the reflectance data.
- **Radiometric Data.** This kind of dataset (such as Arad [S50]) performed radiometric calibration to rectify measurement biases introduced by QE. Radiometric data eliminates the influence of QE but keeps the illumination spectrum, representing the general radiation energy of target scenes.

The spectrum typically quantifies the photon flux (or radiation energy) in each band in optical analysis. In this regard, we implemented radiometric calibration (as detailed in Ref. [S61]) on our acquired HSI dataset to eliminate the

QE bias of raw data. Consequently, the calibrated data is better suited for computational spectral reconstruction tasks, as these data accurately characterize the real radiation energy distribution of target scenes. Therefore, we train the HSI reconstruction model using this calibration dataset. Furthermore, we also utilized a larger training dataset compared to other computational spectral imaging tasks, ensuring both spectral reconstruction accuracy and model generalization as shown in Tab. S3.

5.2 SRNet details

In this section, we describe the overall network architecture and internal module details of SRNet (Spectral Reconstruction Network) in Extended Data Fig. 6a-b. SRNet is a hybrid neural network that combines the core features of Transformer and CNN architectures for efficient, high-precision reconstruction. It employs Unet-shaped architecture as the baseline consisting of an encoder, a bottleneck, and a decoder. The basic component of SRNet is the Spectral Attention Module (SAM), which can focus on extracting the spectral features of HSIs.

First, SRNet exploits a 3×3 convolution layer to transform input images into features $\mathbf{X}_0 \in \mathbb{R}^{H\times W\times C}$ where H, W are spatial dimensions and C is the number of feature channels. The sensing matrix of the HyperspecI sensor also undergoes a 3×3 convolution layer to extract the inherent spectral modulation characteristics, which is fused with $\mathbf{X}_0$ using a 3×3 convolution layer. Second, the fused feature map goes through a down-sample module, a SAM, a down-sample module, a SAM, a down-sample module, a SAM, and a down-sample module to generate hierarchical features. The down-sample module is a stride 4×4 convolution that down-scales the feature maps and doubles the channels. Therefore, the feature maps of the i_{th} stage of the encoder are denoted as $\mathbf{X}_i \in \mathbb{R}^{\frac{H}{2^i}\times\frac{W}{2^i}\times 2^i C}$. Third, $\mathbf{X}_4$ passes through the bottleneck that consists of a SAM. Subsequently, we followed the Unet architecture and designed a symmetrical structure as the decoder. Specifically, the up-sampling module is a stride 2×2 deconvolution layer. Skip connections are used for feature concatenation between the encoder and decoder to alleviate the information loss caused by the down-sampling operations. After passing through the decoder, the feature maps will be added with the initial feature map $\mathbf{X}_0$ and then undergo a 3×3 convolution layer to generate the reconstruction hyperspectral images.

Although attention mechanisms have found widespread application in image reconstruction tasks, they often impose a substantial computational burden. This is particularly evident for self-attention layers in the Transformer framework, which contribute significantly to the overall computational overhead. In many Transformer-based methods, the spatial resolution of input causes a quadratic growth in the time and memory complexity of the key-query dot-production interaction [S62]. We employ transposed query and key attention feature maps in our network to reduce complexity and efficiently capture global context. Instead of performing the dot product operation across spatial dimensions, we perform it across spectral channel dimensions, resulting in linear complexity. This modification allows us to compute cross-covariance across channels and create attention feature maps with implicit knowledge of the spectral information and global context [S52, 62, 63].

From a layer normalized feature map $\mathbf{X}$, we use 1×1 convolutions followed by 3×3 depth-wise convolutions to generate query $\left(\hat{\mathbf{Q}} \in \mathbb{R}^{H \times W \times C}\right)$, key $\left(\hat{\mathbf{K}} \in \mathbb{R}^{H \times W \times C}\right)$ and value $\left(\hat{\mathbf{V}} \in \mathbb{R}^{H \times W \times C}\right)$ projections. Inspired by previous works [S52, 62, 63], We calculate the self-attention by:

$$\text{Atten}\left(\hat{\mathbf{Q}}, \hat{\mathbf{K}}, \hat{\mathbf{V}}\right) = \mathbf{V} \cdot \text{softmax}\left(\sigma \mathbf{K}^{\mathrm{T}} \mathbf{Q}\right). \tag{S4}$$

$$\mathbf{X^*} = W_p \text{Atten}\left(\hat{\mathbf{Q}}, \hat{\mathbf{K}}, \hat{\mathbf{V}}\right) + \mathbf{X} \tag{S5}$$

where $\mathbf{K}^T$ is the transposed matrix of $\mathbf{K}$; W_p is the 1×1 convolution; $\mathbf{Q} \in \mathbb{R}^{\mathbf{HW} \times C}$, $\mathbf{K} \in \mathbb{R}^{\mathbf{HW} \times C}$, $\mathbf{V} \in \mathbb{R}^{\mathbf{HW} \times C}$ are reshaped from original dimension $\mathbb{R}^{H \times W \times C}$. Before using the softmax function, the learnable parameter σ is applied to regulate the magnitude of the dot product of $\mathbf{K}^T$and $\mathbf{Q}$. Additionally, we divide the number of channels into 'heads', and simultaneously learn several attention maps for each head. As shown in Extended Data Fig. 6b, we add a parallel structure to the feed-forward network (FFN) to further extract useful features. First, we apply linear normalization to normalize the feature maps. Then, we involve a 1×1 convolution, GELU [S64], 3×3 depth-wise convolution to capture local information. Next, we use concatenation to fuse two branches of information. Finally, we utilize a GELU activation function and a 1×1 convolution to match the dimension of the input channels.

The training data in this work were primarily acquired using the commercial Figspec-23 hyperspectral camera and GaiaField Pro-N17E-HR hyperspectral camera, both integrated under a push-broom scanning mechanism (see Supplementary Section 5.1). Because the spatial resolution of measurements is very large we randomly divided the calibrated pattern into multiple sub-patterns for training, which also can avoid overfitting to a particular BMSFA encoding pattern. During each iteration, we randomly selected a 512×512 sub-pattern from the original full-resolution BMSFA pattern (2048×2048 for HyperspecI-V1 and 1024×1024 for HyperspecI-V2). This approach can enhance the model's generalization across different spatial positions and encoding patterns. The model is trained with Adam [S65] optimizer ($\beta_1 = 0.9, \beta_2 = 0.999$) for 1×10^6 iterations. The learning rate is initialized to 4×10^{-4} and the Cosine Annealing scheme is adopted. We choose Root Mean Square Error (RMSE), Mean Relative Absolute Error (MRAE), and Total Variation (TV) as the hybrid loss function. we train the model on the Pytorch platform with a single NVIDIA RTX4090 GPU.

5.3 Reconstruction performance comparison

To demonstrate the performance of the model, we have compared the SRNet with other reconstruction algorithms including both model-based and deep learning-based algorithms. The quantitative metrics to evaluate reconstruction precision include Peak Signal-to-Noise Ratio (PSNR), Root Mean Square Error (RMSE), Mean Relative Average Error (MRAE), Structural Similarity Metric (SSIM), and spectral fidelity. The quantitative metrics to evaluate reconstruction efficiency include the number of parameters, floating point operations per second (FLOPs) at two kinds of resolution (2048 $\times$ 2048 pixels and 1024 $\times$ 1024 pixels), and running time per frame image. The specific algorithms for comparison are listed below.

- **Model-based algorithms.** The model-based algorithms typically employ image priors for regularization optimization, such as total variation (TV) [S66], sparsity [S67], non-local similarity (NLS) [S68, 69], and low rank [S70]. The solution is achieved by iterative frameworks such as alternating projection and alternating direction methods of multipliers (ADMM). We utilized three common model-based iterative algorithms for comparison, including the GAP-TV [S71], DeSCI [S70], and PnP [S72] algorithms. Although the model-based algorithms do not require a training process, they involve numerous iterations that result in much running time. Convergence for large-scale images (e.g., 2048 $\times$ 2048) requires heavy computational load, leading to low reconstruction efficiency [S73]. Additionally, the limited representation ability of image priors hinders effective exploration of fine features, leading to relatively low reconstruction accuracy compared to the deep learning-based methods [S74]. As shown in Tab. S4, all of the GAP-TV, DeSCI, and PnP methods take several hours to reconstruct hyperspectral images from only a single frame, which makes them impossible for real-time hyperspectral reconstruction. Moreover, the accuracy of these algorithms is limited especially for large-scale spatial resolution.
- **Deep learning-based methods.** Besides the model-based algorithms, we also conducted a comprehensive comparison among the state-of-the-art (SOTA) deep-learning methods for hyperspectral reconstruction, including HSCNN+ [S75], HRNet [S76], MST++ [S77], HDNet [S78], MST_L [S52], and CST_L [S79]. These models were trained using the same training data as our SRNet (see the Supplementary Section 5.2 for more details), which process was conducted on an NVIDIA RTX4090 GPU and an AMD EPYC 7H12 CPU, with an appropriate batch size on the Pytorch platform. The initial learning rate was set to 4×10^{-4}, with 1×10^{6} iterations using the Cosine Annealing scheduler.

As shown in Tab. S4, the quantitative results were produced on test data with the SNR being 30dB. The results indicate that our SRNet model outperforms others in terms of PSNR, RMSE, MRAE, and SSIM. Besides, we can see that our model shows superior computational efficiency compared to other model-based and deep-learning methods. Specifically, we developed two versions of the SRNet, including SRNet_S and SRNet_L. The SRNet_S has 16 channels of features, and thus contains less parameters than SRNet_L which has 32 channels of features. Consequently, with little sacrifice of reconstruction precision, SRNet_S reconstruction achieves 31 frames and 105 frames at the spatial resolutions of 2048×2048 and 1024×1024, respectively, making it well-suited for real-time reconstruction of high-resolution images. Considering its high efficiency, we employed SRNet_S in other experiments for HyperspecI reconstruction.

Tab S4: Comprehensive performance comparison among different HSI reconstruction algorithms.

	PSNR	RMSE	MRAE	SSIM	Fidelity	Params (M)	FLOPs (T) (2048×2048)	Time (s) (2048×2048)	FLOPs (T) (1024×1024)	Time (s) (1024×1024)
GAP-TV [S71]	21.24	0.0867	2.255	0.617	93.87%	\	\	9164.8	\	2291.2
DeSCI [S70]	21.61	0.0831	2.439	0.653	97.54%	\	\	\	\	238159.1
PnP [S72]	20.07	0.0992	1.921	0.680	95.62%	\	\	9868.8	\	2464.2
HSCNN+ [S75]	38.66	0.0134	0.184	0.970	99.23%	2.140	\	\	2.139	0.23
HRNet [S76]	40.51	0.0111	0.135	0.980	99.61%	7.857	2.674	0.35	0.668	0.083
MST++ [S77]	40.83	0.0105	0.136	0.980	99.60%	1.730	1.615	0.74	0.403	0.18
HDNet [S78]	41.80	0.0095	0.124	0.982	99.63%	2.671	11.200	1.13	2.800	0.28
MST_L [S52]	41.86	0.0094	0.122	0.982	**99.64%**	2.610	2.102	1.45	0.525	0.35
CST_L [S79]	39.62	0.0119	0.146	0.980	**99.64%**	3.910	\	\	0.585	0.39
SRNet_L	**43.98**	**0.0075**	**0.112**	**0.983**	99.62%	1.877	0.384	0.075	0.096	0.016
SRNet_S	42.57	0.0087	0.131	0.979	99.51%	**0.486**	**0.134**	**0.032**	**0.034**	**0.0095**

6 The performance of HyperspcI

The HyperspecI-V1 sensor exhibits a spatial resolution of 2048 × 2048 pixels and a spectral resolution of 2.56 nm. It produces 61 spectral channels at 10 nm intervals over the 400-1000 nm range, with an acquisition speed of 47 frames per second (FPS). We developed the SRNet, a deep learning model based on spectral attention Transformer, to reconstruct hyperspectral images from the acquired raw data, achieving up to 31 FPS using NVIDIA RTX4090 GPU.

The HyperspecI-V2 sensor exhibits a spatial resolution of 1024 × 1024 pixels and a spectral resolution of 8.53 nm. Covering a broad wavelength range from 400 to 1700 nm, it produces 96 spectral bands (61 bands with 10 nm intervals at 400-1000 nm and 35 bands with 20 nm intervals at 1000-1700 nm). It maintains a relatively high acquisition speed of 134 FPS, and a reconstructed speed of 105 FPS using SRNet with NVIDIA RTX4090 GPU. The spectral resolution is lower than that of the HyperspecI-V1 sensor because of the broader wavelength range under a fixed bandwidth [S80].

Tab S5: The performance of HyperspecI-V1 and HyperspecI-V2.

HyperspecI	Chip type	Spatial resolution (pixels)	Acquisition rate (fps)	Reconstruction rate (fps)	Spectral range (nm)	Spectral bands	Spectral resolution (nm)
V1	SONY IMX264	2048 × 2048	47	31	400 - 1000	61	2.56
V2	SONY IMX990	1024 × 1024	124	105	400 - 1700	96	8.53

In the next, we comprehensively evaluate the performance of HyperspecI-V1 and HyperspecI-V2 in detail from multiple aspects, including signal-to-noise ratio, frame rate, noise resistance, dynamic range, light throughput, and thermal stability.

6.1 Signal-to-noise ratio

The spatial-spectral resolution is also affected by the sensor's signal-to-noise ratio (SNR). To quantitatively evaluate the influence, we first employed the EMVA Standard 1288 [S81, 82] to calibrate the sensor's SNR of HyperspecI-V1 and HyperspecI-V2, respectively. Specifically, for a sensor with linear systems, the SNR of the digital output signal is defined as:

$$\mathrm{SNR} = \frac{\mu_{\mathrm{y}} - \mu_{\mathrm{y.dark}}}{\sigma_{\mathrm{y}}}, \tag{S6}$$

where μ_{y} represents the mean of the output signal, $\mu_{\mathrm{y.dark}}$ represents the mean of the dark signal, and σ_y represents the standard deviation of the output signal. We utilized a light source of Thorlabs SLS302 (10W output optical power, 1115 lux) for illumination, and employed a standard reflective whiteboard (barium sulfate) as the target scene. We used the sensors to acquire images with the same exposure time of 1 ms. This setting of illumination intensity and exposure time ensures that the mean of the output signal μ_{y} is around half of the maximum, which is a common setting for SNR calibration [S81]. We repeated this calibration process 100 times, and produced the averaged SNR of HyperspecI-V1 being 36.87 dB (the SNR of the underlying SONY IMX264 chip is 40.2 dB), and that of HyperspecI-V2 being 36.94 dB (the SNR of the underlying SONY IMX990 chip is 52.2 dB).

6.2 Frame rate

The frame rate is a crucial consideration in computational hyperspectral imaging. We analyze the factors that limit HyperspecI's frame rate from two aspects, including hardware acquisition speed and software reconstruction speed.

Hardware acquisition speed. The sensor's hardware acquisition speed is mainly determined by the sensor's light throughput and transmission bandwidth. Light throughput determines a sensor's requisite exposure time under a certain kind of illumination. Especially, in low-light imaging conditions, lower light throughput necessitates longer exposure time. As illustrated in Fig. 2d of the Main Manuscript and Supplementary Section 6.5, the HyperspecI-V1 and HyperspecI-V2 sensors achieve the average light throughput of 71.8% and 74.8%, respectively. Such light throughput is much higher than the existing sensors utilizing narrowband filtering and microstructure spectral modulation. This advantage provides great potential of the reported HyperspecI technique for various low-light applications.

Transmission bandwidth restricts the upper limit of signal acquisition speed. With a fixed transmission bandwidth, a higher spatial resolution results in a lower acquisition speed. The two sensor chips we employed (SONY IMX264 and SONY IMX990) both utilize a common USB3.0 interface, achieving imaging frame rates of 47 frames (2048 $\times$ 2048 pixels) and 134 frames (1024 $\times$ 1024 pixels), respectively. To further enhance imaging speed, higher-quality data acquisition cards can be employed to increase signal transmission bandwidth, which however would increase cost and decrease the integration level.

Software reconstruction speed. The software reconstruction speed is constrained by computational power and the complexity of the reconstruction algorithm. In all the experiments presented in our manuscript, we utilized an NVIDIA RTX4090 GPU (82.58 TFLOPs) for model training and testing. To process high-resolution raw data efficiently, the reported SRNet adopts a lightweight design to acquire more video frames.

We performed a comprehensive reconstruction performance comparison between the SRNet and other hyperspectral reconstruction algorithms, as presented in Tab. S4. To demonstrate reconstruction efficiency, the table also presents the frame rates of different models at 2048 × 2048 pixels and 1024 × 1024 pixels. We can see that the acquisition frame rate of HyperspecI-V1 is 47 frames, with a reconstruction frame rate of 31 frames. For HyperspecI-V2, the acquisition frame rate is 134 frames, with a reconstruction frame rate of 105 frames. The experiments validate that the SRNet model fulfills the real-time requirement for hyperspectral image reconstruction.

6.3 Noise resistance

Tab S6: The noise resistance performance of the reported SRNet.

SNR	PSNR	RMSE	MRAE	SSIM	Fidelity
10	41.84	0.00965	0.1561	0.9687	99.40%
20	43.65	0.00775	0.1193	0.9810	99.58%
30	43.98	0.00749	0.1117	0.9829	99.62%
INF	43.99	0.00747	0.1106	0.9831	99.62%

Next, we further used synthesized measurements to assess HSI reconstruction quality under different SNR conditions (SNR=10, 20, 30, and noiseless), as outlined in Tab. S6. From the results, we can see that as the noise level increases, the reconstruction quality of our HyperspecI sensor keeps above 41 dB (PSNR) even under an input SNR of 10 dB. Considering that the above calibrated SNR is higher than 36, the results validate that our HyperspecI sensor exhibits strong noise resistance, effectively mitigating interference from external environmental noise, and ensuring both spatial resolution and spectral accuracy even in low-light conditions where environmental and system noise degrades the sensor's spatial resolution.

6.4 Dynamic range

We designed and integrated a Broadband Multi-Spectral Filter Array (BMSFA) with the imaging chip. The BMSFA comprises different modulation materials, each having different spectral modulation effects on incident light. Indeed, it has a certain impact on the dynamic range of spectral coupled measurements. Meanwhile, for the computational spectrometers and imaging sensors, the dynamic range is also determined by the reconstruction algorithm and measurement noise. As indicated by the subsequent analysis and calculation, the dynamic range of the HyperspecI sensor is 68.71 dB, and the dynamic range of spectral reconstruction increases as the level of measurement noise decreases. The detailed calculation process is as below.

(1) **The dynamic range of measurements.**

Following the definition of dynamic range [S83], it is calculated as the ratio of the maximum intensity it can measure and the minimum signal that is distinguishable from the noise:

$$DR = \log_{10}\left(\frac{N_{\text{sat}}/\operatorname{Max}_j\left[A\left(\lambda_j\right)\right]}{N_{noi}/\operatorname{Min}_j\left[A\left(\lambda_j\right)\right]}\right), \tag{S7}$$

where λ_j represents the signal of a pixel in the j band, and $A(\lambda_j)$ represents the sensor's throughput that is calculated as

$$A\left(\lambda_j\right) = A_{\text{det}}\,\Omega_{\text{det}}\,T_{\text{opt}}\,(\lambda)\eta_{\text{det}}\,(\lambda), \tag{S8}$$

where A_{det} denotes the photosensitive element area of an individual pixel (25 μm^2 for the employed SONY IMX990 chip), Ω_{det} represents the solid angle of the focused light cone incident on the chip ($\sim 30°$ for the employed SONY IMX990 chip), $T_{opt}(\lambda)$ signifies the optical transmission ratio (exceeds 80% for the employed SONY IMX990 chip), and $\eta_{det}(\lambda)$ indicates the sensor's quantum efficiency (exceeds 75% for the employed SONY IMX990 chip).

The variables N_{sat} and N_{noi} in Eq. (S7) denote the saturation and noise levels of the photoelectrons, which are calculated as

$$\begin{aligned} N_{\text{sat}} &= A\left(\lambda_j\right) tL_{\text{sat},j}, \\ N_{noi} &= A\left(\lambda_j\right) tL_{\text{noi},j}, \end{aligned} \tag{S9}$$

where $A(\lambda_j)$ represents the sensor's throughput as introduced above, t denotes the integration time, $L_{\lambda,\text{sat},j}$ and $L_{\lambda,\text{noi},j}$ represent the radiative intensity of incident photons in the j band corresponding to the saturation and noise levels, respectively.

Utilizing the specific parameters of the SONY IMX990 sensor chip, where A_{det} is 25 μm^2, Ω_{det} is 30°, Ω_{det} exceeds 80%, and $\eta_{det}(\lambda)$ exceeds 75%, the calculation yields a dynamic range of ∼72 dB for the sensor chip without integrating the BMSFA.

Subsequently, we employed the HyperspeI sensor equipped with BMSFA to measure the full-spectrum responses of the ten organic dyes and six nano-metal oxide materials, to update the values of $\operatorname{Max}_j\left[A\left(\lambda_j\right)\right]$ and $\operatorname{Min}_j\left[A\left(\lambda_j\right)\right]$ for each material. Then, we selected the highest maximum throughput as V_{max} and the lowest minimum throughput as V_{min} (as presented in Tab. S7). With these two variables, the updated dynamic range calculation is denoted as

$$DR_{\text{update}} = \log_{10}\left(\frac{N_{\text{sat}}\,/V_{\text{max}}}{N_{noi}/V_{\text{min}}}\right). \tag{S10}$$

Following the formula, the updated dynamic range of the HyperspecI sensor is 68.71 dB. Compared to the original dynamic range of ∼72 dB, the calculation result indicates a slight reduction (3.29 dB) in the sensor's dynamic range.

Tab S7: V_{max} and V_{min} of the ten organic dyes and six nano-metal oxide materials used for BMSFA fabrication. These two variables were applied to calculate the dynamic range of the HyperspecI sensor following Eq. (S10).

Sample	$\mathrm{Max_j[A(\lambda_j)]}$	$\mathrm{Min_j[A(\lambda_j)]}$
ODM-1	90.1364	1.6236
ODM-2	92.1512	5.0945
ODM-3	93.0745	1.4631
ODM-4	93.3982	7.2340
ODM-5	92.3322	5.6478
ODM-6	93.2966	3.1851
ODM-7	92.7234	0.0483
ODM-8	92.9978	15.6535
ODM-9	89.9630	34.1318
ODM-10	89.7190	44.6237
NMO-1	76.1518	61.2237
NMO-2	59.3825	0.0818
NMO-3	90.0000	26.6484
NMO-4	68.8982	29.2056
NMO-5	38.6110	26.3048
NMO-6	59.2514	29.0367
V_{max}/V_{min}	93.3982	0.0483

Following the above calculation process, we also calculated the dynamic range changes of both the RGB camera (HIKVISION MV-CA050-12UC) and the MSFA camera (Silios CMS-C), as shown in Tab. S8. We can see that for both the RGB camera and MSFA camera, there exist a reduction in dynamic range (1.27 dB and 3.41 dB, respectively) after integrating the Bayer pattern and the MSFA.

Tab S8: Dynamic range reduction of the BMSFA, RGB, and MSFA sensors after integration of a filtering array.

Architecture	**Exemplar model**	**Chip model**	**Dynamic range reduction**
BMSFA	HyperspecI	SONY IMX990	3.29 dB
RGB	HIKVISION MV-CA050-12UC	SONY IMX264	1.27 dB
MSFA	Silios CMS-C	No disclosure	3.41 dB

(2) **The dynamic range of spectral reconstruction.**

Furthermore, as claimed in Ref. [S84] for computational spectrometers and imaging sensors, the dynamic range is not only determined by the maximum signal (the maximum range of the analogue-to-digital converter (ADC) of the sensor chip) and the minimum signal (the standard deviation of the dark signal), but also by the reconstruction algorithm and measurement noise. In this regard, following the calculation strategy in Ref. [S84], we further studied how the 4 levels of measurement noise (standard deviation $\sigma = 0.0001$, 0.001, 0.01, 0.1) affect the dynamic range of our HyperspecI sensor. Considering a 12-bit ADC (which can produce an integer from 0 to $2^{12} - 1 = 4095$) in the spectral range of 400-1700 nm (96 data points over the 1300 nm range), we set an incident spectrum $\Phi(\lambda)$ with a maximum intensity of 4095 as the target spectrum. The synthesized measurement I_i was calculated as

$$M_i = (1 + \varepsilon) \sum_{\lambda} \Phi(\lambda) T_i(\lambda), \tag{S11}$$

where ε denotes measurement noise with its standard deviation $\sigma = 0.0001$, 0.001, 0.01, 0.1, representing different noise levels. $T_i(\lambda)$ is the transmission spectrum of the i_{th} BMSFA filter. These measurements M_i were input into the reconstruction algorithm to produce the simulated reconstructed spectrum, from which we extracted I_{max} as the maximum intensity in the reconstructed spectrum. This process was repeated 100 times and averaged to produce $I_{max,mean}$.

Next, we calculated the standard deviation $\sigma_{dark,mean}$ of dark signal. We first synthesized the dark counts $M_{dark,i} =\mid 4095(\sigma + \varepsilon_{\text{dark}}) \mid$, where $\varepsilon_{\text{rand}}$ follows a normal distribution with a standard deviation of $\sigma = 0.0001$, 0.001, 0.01, 0.1, representing different noise levels. The dark counts were input into the reconstruction algorithm for dark count spectral reconstruction. Then, we calculated σ_{dark} as the standard deviation of the reconstructed spectral intensities across the entire wavelength range. This procedure was similarly repeated 100 times and averaged to produce $\sigma_{dark,mean}$.

Finally, the dynamic range at each noise level is calculated as $DR = 20\log_{10}((I_{max,mean}/\sigma_{dark,mean})$. Following the above method, we calculated the dynamic range of spectral reconstruction for our HyperspecI sensor, as presented in Tab. S9. From the results, we can see that the dynamic range of spectral reconstruction increases as the level of measurement errors decreases.

6.5 Light throughput

We explored and compared the high light throughput of different hyperspectral imaging techniques. First of all, we would like to illustrate that all of the related techniques for throughput comparison are based on a similar working principle that superposes spectral modulation on a basic monochromatic chip. In this regard, we did not take the underlying chip's quantum efficiency into account for throughput comparison, because different techniques

Tab S9: Simulation on the dynamic range of spectral reconstruction for our HyperspecI sensor at different measurement noise levels. The dynamic range increases as the measurement noise decreases.

Noise Level / σ	**0.0001**	**0.001**	**0.01**	**0.1**
$I_{max,mean}$	4090	4089	4075	3850
$\sigma_{dark,mean}$	1.8	5.3	11.2	89.4
DR (dB)	67.12	57.74	51.19	32.66

employed different chips for experiment demonstration. Instead, we only compared the light throughput of spectral modulation to fairly demonstrate the high-throughput advantage of the reported HyperspecI technique. The specific throughput calculation process for each technique is as follows.

1. Scanning hyperspectral camera: Scanning hyperspectral cameras are typically composed of a diffractive grating, a slit, and a mechanical translation module [S85]. It acquires HSI information by one column of pixels at a time. The entire hyperspectral image is obtained by mechanically scanning in the row dimension. The spectral separation strategy using a slit and grating constrains its light throughput to a large extent. Taking the push-broom FigSpec-23 hyperspectral camera [S40] as an example, the light throughput is calculated considering both the slit loss and the grating loss.

• **Slit Loss:** The FigSpec-23 hyperspectral camera has a slit width of 25 µm. It captures spectral data for only one column of pixels per frame. The mechanical module collects spectral data for 100 columns of pixels per second. To acquire hyperspectral data with a spatial resolution of 960×1230 pixels that is similar to HyperspecI-V2, the FigSpec-23 camera needs at least 10 seconds.

• **Grating Loss:** The transmission grating for spectral separation maintains a transmission efficiency of ∼60%, further reducing the camera's light throughput. The grating spectral separation strategy results in additional decay of the photon flux on each spectral channel, with each channel representing only a fraction of the total incident light energy.

The frame rate of the underlying sensor chip (acA1920-155um) used in the FigSpec-23 hyperspectral camera is 100 FPS @ 960 × 1230 pixels, and it can only capture spectral information for one row of pixels per frame. As a result, the slit reduces the throughput to 0.1% of the incident light. The grating further reduces the throughput to 60%. If the grating split incident light to multiple channels at 10 nm intervals, the calculated light throughput for each channel would be one-ten-thousandth of the total incident light. Thus, the push-broom scanning hyperspectral cameras commonly consume a long time to obtain a hyperspectral image of the target scene, which is unable to acquire spectral information of dynamic scenes.

2. RGB color camera: RGB color cameras utilize Bayer filters (red, green, and blue) to capture raw data, which is subsequently processed to produce color images using interpolation algorithms. Each of the RGB color filters exhibits light transmission within a relatively narrow spectral range. Given that the entire detection range is 400-700 nm, we calculated the integral of the transmission spectra, producing a light throughput of less than 30%.

3. MSFA-based multispectral camera: The MSFA-based multispectral camera (Silios CMS-C [S86]) employs an array of narrowband filters for spectral acquisition, leading to a significant reduction in light throughput. We calculated the integral of each transmission spectrum from 400 nm to 750 nm, producing the camera's light throughput of less than 10%.

4. Coded Aperture Snapshot Spectral Imager (CASSI): The CASSI system utilizes a coded mask (grayscale aperture [S87] or color aperture [S88]) and a 2D dispersive element (prism [S89] or grating [S90]) for spectral encoding, resulting in spectral overlap with photon flux attenuation. The coded mask typically decreases light throughput by more than 50%. Additionally, the introduced relay lens and prism further decrease light throughput.

5. Metasurface based camera: Metasurface is a kind of micro-nano filter, in which diverse spectral responses of freeform-shaped patterns result from the complicated scattering and coupling of the propagative Bloch modes in the metasurface layer [S91–93]. We estimated its light throughput through spectral integral over the working spectral range, which is averaged on the multiple filters and produces less than 50%.

6. Fabry–Pérot filter based camera: The Fabry–Pérot filter also exhibits spectral modulation characteristics, and can be integrated with sensor chips [S56]. This filter selectively transmits or reflects specific wavelengths of light based on resonance conditions through multiple interference and reflection within a cavity formed by parallel high-reflectivity mirrors. As reported by Ref. [S56], the light throughput of the Fabry–Pérot filters is 45%.

7. Our HyperspecI sensor: As detailed in the Main Manuscript, we integrated the broadband multispectral filtering array on the monochrome chip to produce the HyperspecI sensor. As shown in Fig. 2d, we calculated the light throughput of HyperspecI sensors across each spectral band. Similar to the RGB color camera and the MSFA-based multispectral camera, we calculated the integral of each transmission spectrum, producing the light throughput of HyperspecI-V1 and HyperspecI-V2 up to 71.8% and 74.8%, respectively.

6.6 Statistical evaluation

We statistically evaluated the reconstruction quality on a testing dataset including 100 indoor and outdoor scenes, in which the scenes are different from the training dataset. The specific dataset construction details are referred to Supplementary Section 5.1.

For quantitative metrics, we utilized the spectral fidelity and Root Mean Square Error (RMSE) to evaluate spectral reconstruction accuracy for the entire spectral range and each band, respectively. Spectral fidelity describes reconstruction accuracy by correlating reconstructed spectra with ground-truth ones, with smaller values indicating higher accuracy. As presented in the Supplementary Section 5.3, our experiments resulted in an average spectral fidelity of 99.62%. The RMSE metric assesses HSI reconstruction error for each band over all the spatial locations and different scenes.

As for the repeatability of reconstruction performance, the thermal stability study in the Supplementary Section 6.7 may help for validation. In the study, we evaluated the reconstruction stability at different times (30-minute intervals for each temperature) and different temperatures (20°C, 40°C, 50°C, 60°C, 70°C). The results are shown in Extended Data Fig. 4d and f. We employed the minimum Pearson's correlation coefficient (minPC) to quantitatively evaluate reconstruction repeatability under different circumstances. The minPC metric presents the minimum Pearson's Correlation coefficient [S94] among the spectral curves at different times and temperatures. We can see that the minPC of reconstructed spectra in certain conditions reaches 0.99, which demonstrates the repeatability of the reconstructed results.

6.7 Thermal stability

The impact of temperature on the photoresist and modulation mask is a crucial factor to be considered in HyperspecI applications.

First of all, we would like to illustrate that the SU-8 2000, as a negative photoresist for near-ultraviolet (UV) lithography, is particularly suitable for modulation mask fabrication, owing to its nature as a thermosetting polymer [S95]. Being an aromatic compound, it not only boasts high flexibility but also possesses exceptional thermal and chemical stability [S96]. These attributes ensure that SU-8 maintains structural and chemical integrity during the UV exposure and photolithography process, which is crucial for the accuracy and repeatability of photolithographic patterning. In the photolithography process, the post-exposure bake and hard bake are with temperatures of 95°C and 150°C, respectively [S97]. These specified temperatures are imperative for ensuring the complete cross-linking of the SU-8 photoresist. This cross-linking process is fundamental in achieving the structural integrity and stability required for the photoresist pattern. Upon ultraviolet (UV), the cured areas of SU-8 photoresist generally exhibit increased thermal stability [S98]. The curing and cross-linking lead to a higher melting point in the exposed regions, requiring a higher temperature to melt when heated. For film mechanical properties, the softening point is 210°C, and the critical temperature for thermal stability in air is 279°C. Based on the heat dissipation capacity of the camera itself, experimental tests have shown that the temperature of the device within the working temperature range (0-50°C [S99]) does not exceed 70°C. Therefore, the heat generated by common operation will not affect the stability of the photoresist and modulation mask.

Next, we implemented two additional experiments to evaluate the thermal stability of the BMSFA modulation mask and the integrated HyperspecI sensor.

1. Thermal stability of the BMSFA modulation mask

We built an additional experiment setup for the thermal stability study. The setup is based on the Kohler illumination principle [S100] as depicted in Extended Data Fig. 4a, which consists of a light source (Thorlabs SLS302 with a stabilized quartz Tungsten-Halogen lamp of 10W output optical power), an illumination module (composed of optical lens, aperture stop, field stop, beam splitter of Thorlabs VDFW5/M, and objective lens of Olympus microscope objective A 10 PL 10×0.25), a heating stage (JF-956, 30-400°C), and several support components (Thorlabs CEA1400) and a fine-tune module (GCM-VC 13M). The fabricated BMSFA modulation mask was placed on the heating stage, and the temperature was sequentially increased with a step of 10°C from 20°C (room temperature) to 200°C. After the temperature stabilized for each step (waiting for 10 minutes after the actual temperature reached the set temperature), we collected an image of the mask, as shown in Extended Data Fig. 4b. From the results, we can see that after heating, the structure of the modulation mask did not show significant changes.

2. Thermal stability of the HyperspecI sensor

In this experiment, we placed the HyperspecI sensor on the heating stage, and employed it to acquire hyperspectral images of the same scene at different operation temperatures, with subsequent comparisons regarding image similarity and spectral consistency. According to the manual sheet provided by SONY company, the sensor chip's operational temperature spans from 0 to 50°C, with the common operating surface temperature being around 37°C at room temperature (20 °C). To assess the sensor's reconstruction performance at varying temperatures, the sensor was fixed on the heating stage and incrementally heated from 40°C to 70°C with 10°C intervals. At each temperature, the sensor was powered on for 1 hour to achieve temperature stabilization. Extended Data Fig. 4d-f illustrates the hyperspectral imaging performance of the same scene at different temperatures.

Specifically, the experimental scenario is depicted in Extended Data Fig. 4c and e depicts the acquired raw data obtained at different temperatures. Extended Data Fig. 4e presents synthesized RGB images of reconstructed hyperspectral data at varying temperatures, and Extended Data Fig. 4f displays the average spectral curves of different regions in the reconstructed hyperspectral data. The minimum Pearson correlation coefficient for the same area at different temperatures was 0.99, indicating a negligible impact of temperature variations within the normal range on the HyperspecI sensor. Also, we further employed SSIM and PSNR to quantitatively assess the effect of temperature variations on spectral image reconstruction (Extended Data Fig. 4d). These two metrics were calculated on each two hyperspectral raw data for similarity evaluation. The minimum value of 0.9715 for SSIM and 38.25 dB for PSNR reveals a satisfactory consistency among different temperatures.

To sum up, the above experiments validate that within the normal operation temperature range, the HyperspecI sensor's performance exhibits little susceptibility to variations in temperature.

7 Intelligent agriculture applications

7.1 Experiment prototype

We have conducted comprehensive research on two applications and designed an integrated device capable of non-contact, rapid, and high-precision chlorophyll detection, as well as soluble solid content (SSC) detection. Fig. 4a illustrates the development of a hyperspectral imaging system comprising our HyperspecI sensor, an opaque dark box, a scattering plate, and a bimodal LED/broad spectrum light source. This system operates in two modes:

1. Transmission spectra acquisition mode: In this model, the light source was comprised of several LEDs. The resulting light passes through the leaf and onto the scattering plate area. The acquired transmission spectra are processed using the model to calculate the leaf SPAD value.

2. Reflectance spectra acquisition mode: In this model, an apple is placed on the scattering plate position, and the LED is replaced with a broadband light source (LED light sources can also be used after calibration). HyperspecI sensor is used to obtain the reflectance spectra of the apple, which are then fed into the model for SSC prediction.

7.2 SSC prediction of apples

Apple is a widespread agricultural product with considerable nutritional and commercial value. Currently, people pay more attention to the internal quality of apples, where SSC is one of the most crucial indicators that determines taste and influences consumer purchasing decisions. Traditional methods like Brix refractometry are destructive, time-consuming, and limited to a small number of samples. Near-infrared (NIR) spectroscopy can be used as a rapid and non-destructive method to characterize the internal attributes of horticultural products. The propagation of NIR spectra through fruit tissue is affected by the vibration frequencies of hydrogen-containing groups associated with their chemical composition and microstructure. The detection procedure for SSC in apples using HyperspecI sensor can be summarized as follows:

(1) Selecting enough apples as a train/calibration set for modeling analysis;

(2) Obtaining the spectral information of the selected apples;

(3) Employing traditional destructive testing to determine the SSC in the regions of interest (ROIs) in the apples;

(4) Preprocessing the original spectral data and establishing the mathematical model;

(5) Evaluating the effectiveness of the model using an independent test/prediction set.

1. Sample preparation

A total of 54 “Fuji” apples without physiological decay or physical defects were obtained from a local retailer in Beijing, China. All samples were washed, numbered, and stored in a laboratory for 24 hours to reach room temperature and reduce measurement variations caused by temperature changes.

2. Acquisition of reflectance spectra and SSC values

The hyperspectral reflectance imaging system was established, depicted in Fig. 4c. It consists of a HyperspecI sensor, a broadband light source (Thorlabs, SLS302) covering the range of 360-2500 nm, a collimating lens, and optical supports. Marking and measuring were conducted at four equidistant regions of interest (ROIs, 3 cm $\times$ 3 cm) distributed along the equatorial region of each apple with each ROI considered as an independent sample. After removing a bruise sample, a total of 215 spectra were obtained. SSC measurements were immediately carried out using the traditional destructive method after spectra acquisition. A piece of flesh with a thickness of about 10 mm was extracted from the equatorial position, and apple juice was squeezed onto a digital refractometer (DiFluid Air, DiFluid, China) to determine the SSC value. Subsequently, 163 and 52 samples were randomly selected as a train (calibration) set and a testing (prediction) set, respectively, with a proportion of 3:1. These sets were utilized to develop and test the calibration model and evaluate the usability of the reconstructed spectra. It is important to note that there were no duplicate samples in the train or test set, as shown in Fig. 4c.

3. Model establishment and evaluation

(1) Spectral preprocessing

To improve the performance of the model, preprocessing of the raw spectra (S) is necessary. The white reference (W) was obtained from a standard white-board with a reflectance of approximately 99%, and the dark reference (D) was captured by turning off the lamp and fully covering the optical lens. The corrected spectra (R) were calculated according to the equation:

$$R = \frac{S - D}{W - D} \tag{S12}$$

(2) Partial least squares (PLS) regression

Partial least square (PLS) regression is one of the most effective multivariate statistical methods widely employed in NIR analysis due to its insensitivity to collinear variables and large variable quantities. The underlying principle of PLS involves constructing a linear regression model with latent variables (LVs) by projecting the original X variables (spectral data) and y variables (SSC) onto an orthogonal space, thereby maximizing the covariance between LVs. In this research, the optimal number of LVs was determined through the 10-fold cross-validation of calibration samples, when the root mean square error of cross-validation (RMSECV) reached the minimum. Subsequently, the calibration model was tested to predict the SSC of the independent prediction set. The model can be expressed as:

$$y = bX + e \tag{S13}$$

where b and e represent regression coefficients and model offset, respectively. The performance of PLS model was evaluated using the correlation coefficient

(R) and the root mean square error (RMSE).

$$R_c, R_p = \sqrt{\sum_{i=1}^{n}(y_{pi} - y_{ai})^2} \Bigg/ \sqrt{\sum_{i=1}^{n}(y_{pi} - y_{mean})^2} \tag{S14}$$

$$RMSEC, RMSEP = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_{pi} - y_{ai})^2} \tag{S15}$$

where y_{pi}, y_{ai} and y_{mean} represent the predicted value, the actual value, and the actual mean value of the i_{th} sample, respectively. n is the number of corresponding data set. Generally, satisfactory models should possess higher values of R_c and R_p, and lower values of RMSEC and RMSEP.

(3) Experiment results

The spectral curves exhibit prominent peaks near 640 nm and 810 nm, as well as two troughs near 680 nm and 970 nm. The peak at 640 nm is primarily related to the color of the apple's epidermis [S101], while the trough near 680 nm should be relevant to the chlorophyll represent the color feature [S25]. The peak near 810 nm probably contains information related to the third overtone of the band N-H [S102], and the trough near 970 nm is associated with the combination effect of second overtone of band O-H and water [S26, 27]. These characteristics highlight the potential of the HyperspecI sensor for analyzing SSC distribution in apples. PLS model was utilized to extract potential information from complex spectra. The optimal number of LVs is determined by the lowest value of RMSECV to avoid poor prediction accuracy caused by few LVs and overfitting caused by more LVs. Fig. 4c presents the scatter plot of measured and predicted SSC, where the points closer to the solid line indicate better model performance. For the full-spectrum-PLS model, RMSECV has the lowest value when 17 LVs were used. The correction coefficient and RMSE were 0.8264 and 0.6132% for the training set, and 0.6162 and 0.7877% for the test set, respectively. The relative error of the prediction set amounts to only 5.30%. Fig. 4d presents the RGB images and reconstructed hyperspectral images of leaves and apples. Additionally, the random selection of pixels in the images allows a comparison between the ground-truth spectrum and corresponding reconstruction, which shows a small difference that indicates high spectral accuracy.

8 Real-time health monitoring applications

8.1 Blood oxygen detection

We obtained pulse waveform signals (PPG waveform) using the photoplethysmography (PPG) method, which is reflected or transmitted through the process of blood tissue organ transmission (Extended Data Fig. 5a-c). The absorption spectra of oxygenated hemoglobin (HbO_2) and deoxygenated hemoglobin (Hb) exhibit characteristic differences in human blood tissue.

Exploiting this characteristic, we can calculate blood oxygen saturation (SpO_2). In the wavelength range of 600-1000 nm, incident light effectively transmits through human tissue, with minimal absorption by water but significant absorption by deoxygenated and oxygenated hemoglobin, which have different absorption levels. Traditional pulse oximeters measure SpO_2 by exploiting the differential light absorption of HbO_2 and Hb at two specific wavelengths. Based on the absorption spectra of HbO_2 and Hb shown in the figure, we can select two characteristic wavelengths, $\lambda_1 = 780$ nm and $\lambda_2 = 830$ nm. The absorbance by HbO_2 is higher at λ_2 than at λ_1, whereas the absorbance by Hb is higher at λ_1 than at λ_2.

The Lambert-Beer law, widely employed for measuring solute concentration based on optical transmission, states that absorbance is directly proportional to concentration. The pulse oximeter assumes that the pulsatile component (AC component) of optical absorption originates from pulsating arterial blood, while the non-pulsatile component (DC component) includes contributions from non-pulsating arterial blood, venous blood, and other tissues. The pulsatile signal (AC) can be normalized by the non-pulsatile signal (DC) at λ_1 and λ_2. By obtaining the time-domain signals of light transmitted through the fingertip at these two wavelengths, we can calculate the AC and DC components at these spectral bands. Through calibration, the following formula can be derived to measure the desired SpO_2:

$$SpO_2 = k_1 * R^3 - k_2 * R^2 + k_3 * R + C \tag{S16}$$

where $R = \frac{I_{AC,\lambda_1}/I_{DC,\lambda_1}}{I_{AC,\lambda_2}/I_{DC,\lambda_2}}$, $k_1 = 10.0002$, $k_2 = -52.887$, $k_3 = 26.871$, $C = 98.283$.

8.2 Material diffusion examination

In this experiment, we conducted an effluent diffusion monitoring study to explore its capability for water quality detection. We simulated water of varying quality by preparing solutions with different materials and observed their diffusion processes to mimic the water quality monitoring process. We prepared two solutions with similar colors but different components: one with standard magenta dye solution #2, and another with the same concentration of magenta dye combined with infrared absorption material solution #1 (in a 20:1 ratio). The experimental results demonstrate the real-time detection capabilities of the HyperspecI sensor, enabling precise acquisition of spatial-spectral information in dynamic scenes.

This experiment also aims to validate the effectiveness of the additional spectral information collected by our HyperspecI sensor in distinguishing materials. Therefore, we did not use advanced deep-learning-based segmentation models that incorporate various prior assumptions, but instead used the classic machine-learning-based k-means clustering algorithm [S103] for image segmentation of RGB and hyperspectral images. The adopted k-means clustering algorithm includes the following three steps:

(1) We first loaded the RGB video and the hyperspectral video frame by frame, and selected the sector area of diffusion in each frame as the position to be segmented. Each RGB image has 3 channels, and each hyperspectral image contains 61 channels at 10 nm intervals in the range of 400-1000 nm.

(2) We set the cluster number K = 3, and the cluster center was randomly initialized. The petri dish contains two different materials but with similar colors.

(3) For each pixel location $p(x, y)$, we calculated its Euclidean distance from the cluster center c_k:

$$d = \mathbf{p}(x, y) - \mathbf{c}_k. \tag{S17}$$

Each pixel location of the RGB image contains 3-channel information, and that of the hyperspectral image contains 61-channel information. According to the distance d, all the pixels were assigned to the nearest center point. Then, we recalculated the new cluster center position following

$$c_k = \frac{1}{k} \sum_{y \in C_k} \sum_{x \in C_k} p(x, y). \tag{S18}$$

After repeating this process until convergence, we reshaped the clustered pixels onto corresponding images. If the segmentation process is effective, it can accurately display the diffusion process of the two materials in the aqueous solution and the clear solution boundary. The segmentation results, as presented in Fig. 5 of the main manuscript, indicate that the images captured by the RGB camera did not accurately segment the diffusion of the solutions. In contrast, the images captured by the HyperspecI camera successfully distinguished and labeled the diffusion processes of the two solutions.

9 Near-infrared applications

To validate the performance of our HyperspecI sensor in the NIR range, we conducted experiments on textile classification and apple bruise detection.

9.1 Textile classification

The textile industry plays a crucial role in our daily life. Various textile products such as personal clothes and household products are indispensable. With the exponential growth of the textile industry, the direct disposal and underutilization of waste textiles put a strain on resources and caused significant environmental issues [S33]. Efficient recovery and reuse of these materials are crucial for environmental sustainability [S34]. Currently, textile classification, a key process in this context, is predominantly manual with limited automation, resulting in poor accuracy and efficiency in recognition [S104]. Together with chemometric modeling, near-infrared measuring offers a potential solution to automatically classify, identify, and quantify substances [S33].

In this experiment, we prepared 204 samples including different cotton and polyester fabrics. The specific composition of each sample was provided by the corresponding manufacturer. Given the diverse appearance of these samples, their classification by visual inspection is challenging. Therefore, we first acquired the NIR spectrum of each sample using our HyperspecI sensor, as shown in Fig. 6a. Then, 150 samples (75 cotton and 75 polyester) and 54 samples (27 cotton and 27 polyester) were selected through the Kennard-Stone algorithm [S105] as the train set and the testing set, respectively. Subsequently, the support vector machine (SVM) algorithm [S106] was employed for the qualitative classification of waste textiles to facilitate the automatic identification of fabric categories [S34]. For the testing set, the overall classification accuracy was 98.15%, which is comparable to Ref. [S107] for pure fabrics identification. The experiment validates that the HyperspecI sensor is applicable for accurate pure textile-waste classification.

9.2 Apple bruise detection

Computer vision-based detection of fruit size, shape, and color features has been widely applied [S108]. However, the detection of fruit skin defects mainly relies on manual identification. Especially, the hidden defects (e.g., minor bruises, early rotting) among the epidermal defects may further rot and spread to other fruits over time, causing economic losses. Therefore, it is essential to detect the fruits that carry hidden defects [S109]. Manual sorting of hidden defects requires a lot of repetitive labor and is susceptible to subjective factors with poor consistency. Further, the hidden defects caused by mechanical collision usually occur under the surface of the skin, and their appearance is similar to that of normal tissues, which is difficult to be recognized by naked eyes [S110]. Due to the absorption of near-infrared light by water, there is a significant difference between the spectral curves of bruised and normal areas in the near-infrared wavelength band. Notable characteristics are observed at wavelengths of 1060 nm, 1260 nm, and 1440 nm [S111].

In this experiment, we prepared 56 'Qixia' Fuji apples for demonstration. Each quadrant of an apple, demarcated by rotating around its stem-to-calyx axis in 90-degree increments, was individually numbered and classified as a separate sample, amounting to a total of 224 samples. Next, a 30 g solid metal ball was dropped through a 30 cm steel pipe to systematically create bruises on random locations of each apple side. Then, we employed our HypersepcI sensor and an RGB camera (for comparison) to respectively acquire hyperspectral and color images of these apples, constructing two image datasets. Each dataset consists of 224 images, 184 images of which were applied to train a YOLOv5-based image detection model [S112], and the rest 40 samples were input to the model for testing. For the hyperspectral images, following

Ref. [S111], we chose 1060 nm, 1260 nm, and 1440 nm as pivotal wavelengths for bruise identification. We processed the spectral images to create synthesized color representations, distinctly marking the bruised regions for enhanced visualization.

The results are presented in Extended Data Fig. 6d-g, which reveals a significant contrast in detection accuracy between visible light and NIR light. We used the Precision (P), Recall (R), mAP50, and mAP50-95 metrics [S112] for quantitative evaluations. From the results, we can see that all the indicators of apple bruise detection based on RGB images are much lower than that based on NIR images. This dues to the insensitivity of the visible part to the bruised area and the difficulty of labeling by manual methods. NIR images exhibit superior precision in identifying bruised areas on apples, demonstrating the potential of our device for efficient NIR detection tasks in intelligent agriculture.

10 Remote detection application

With the irreplaceable advantage of high signal-to-noise ratio (SNR), high resolution, and dynamic hyperspectral imaging capabilities, our HyperspecI possesses apparent advantages over existing methods in detecting small or distant targets within low-light conditions, targeting dynamic scenes, among others (as partially validated in Fig. 3 of the main manuscript). Additionally, the compact size, lightweight, and high integration level of HyperspecI make it suitable for deployment on platforms with limited payload capacity.

To highlight the high SNR and dynamic hyperspectral imaging features of HyperspecI, we demonstrate its adoption in the astrophotonic application for remote detection. We chose a cloudy, foggy night with relatively moonlight to demonstrate a low-light environment. As shown in Extended Data Fig. 1a, we employed a telescope (CELESTRON NEXSTAR 127SLT, 1500 mm focal length, 127 mm aperture Maksutov-Cassegrain) to image the moon, and compared the imaging results of our HyperspecI sensor with that of line-scanning hyperspectral camera (FigSpec-23) and mosaic multispectral camera (Silios CMS-C). The target scenes include the Mare Crisium (see #2 in Extended Fig. 1c) and Mare Fecunditatis (see #1 and #4 in Extended Data Fig. 1c) regions of the moon during the crescent phase. Our HyperspecI's acquisition frame rate is set to 47 fps, with an exposure time of 21 ms. The mosaic multispectral camera has an acquisition frame rate of 30 fps, with an exposure time of 33 ms. The line-scanning hyperspectral camera requires approximately 100 seconds to capture one HSI frame. We can see that the results of the HyperspecI sensor present fine details of lunar topography, and the reconstructed spectrum corresponds well with the ground truth (publically provided at http://www.olino.org/blog/us/articles/2015/10/05/spectrum-of-moon-light#comment-3920724). In contrast, the results of the mosaic multispectral camera contain serious measurement noise due to limited light throughput, and the topography details are buried. The results of the line-scanning hyperspectral camera suffer from a similar degradation, and severe scanning overlapping exists since the moon was moving during the line-scanning process. This experiment demonstrates the unique high-light throughput advantage of our HyperspecI sensor, which leads to high imaging SNR that enables the acquisition of dynamic, remote, and fine details in low-light conditions. More dynamic imaging results are presented in Extended Data Fig. 1d. From the above experiment, we can see that our HyperspecI exhibits irreplaceable advantages in hyperspectral imaging in dynamic or low-light environments, which is impossible to achieve with other existing methods.

In summary, the high SNR, high temporal-spatial-spectral resolution, and wide working wavelength range of HyperspecI break through the limitations of existing technologies in multiple fields, making high-precision hyperspectral imaging possible in various complex scenarios.

11 Metamerism Experiment

In the visible band, a phenomenon called metamerism refers to different spectra projected to the same color. We further implemented an additional metamerism experiment and validated our sensor's ability to distinguish between items with the same colors but different spectra, thus reaffirming the sensor's strong generalization performance. The experiments' details are as follows. As shown in Extended Data Fig. 2a, we tested the real and fake potted plants, marking points with the same color in a(ii). We marked the points of potted plants with the same color (by comparing their 3-channel values), in which the red points on the true potted plant and the yellow points on the fake potted plant have the same color. Extended Data Fig. 2a(iii) shows the original measurement of our HyperspecI sensor, and the reconstructed hyperspectral image is shown in a(v). We plotted the spectra of points **P1** and **P2** with the same color on both real and fake plants, as shown in Extended Data Fig. 2a(iv). It's evident from a(iv) that the leaves of the real plant exhibit distinct spectral features due to variations in chlorophyll and water content (as shown in the blue block of the a(iv)), whereas the fake plant displays completely different discernible spectra. Furthermore, we tested real and fake strawberries, which have nearly identical appearances, textures, and colors, as shown in Extended Data Fig. 2b. Extracting the spectra of points **P1** and **P2** with the same color in both real and fake strawberries, we observed that the real strawberries exhibit distinct absorption peaks at 670 nm and 750 nm due to chlorophyll and anthocyanins content, while the spectra of fake strawberry appear smoother. These experiments not only validate the high-resolution and high-precision hyperspectral imaging capabilities of our HyperspecI technique, but also demonstrate its ability to effectively differentiate materials with the same color but different spectra.

References

S1 Antaris, A.L., Robinson, J.T., Yaghi, O.K., Hong, G., Diao, S., Luong, R., Dai, H.: Ultra-low doses of chirality sorted (6, 5) carbon nanotubes for simultaneous tumor imaging and photothermal therapy. ACS nano **7**(4), 3644–3652 (2013)

S2 Bai, J., Liu, Y., Jiang, X.: Multifunctional PEG-GO/CuS nanocomposites for near-infrared chemo-photothermal therapy. Biomaterials **35**(22), 5805–5813 (2014)

S3 Kramer, I.J., Sargent, E.H.: The architecture of colloidal quantum dot solar cells: materials to devices. Chemical reviews **114**(1), 863–882 (2014)

S4 He, Y.Q., Liu, S.P., Kong, L., Liu, Z.F.: A study on the sizes and concentrations of gold nanoparticles by spectra of absorption, resonance Rayleigh scattering and resonance non-linear scattering. Spectrochim. Acta. A **61**(13-14), 2861–2866 (2005)

S5 Pirker, L., Višić, B., Kovač, J., Škapin, S.D., Remškar, M.: Synthesis and characterization of tungsten suboxide W_nO_{3n-1} nanotiles. Nanomaterials **11**(8), 1985 (2021)

S6 Sreekumar, S., Joseph, A., Kumar, C.S., Thomas, S.: Investigation on influence of antimony tin oxide/silver nanofluid on direct absorption parabolic solar collector. J. Clean. Prod. **249**, 119378 (2020)

S7 Shen, B., Wang, Y., Lu, L., Yang, H.: Enhanced spectral modulation of Cs_xWO_3 nanocrystals through anionic doping for energy-efficient glazing. Sol. Energy Mater. Sol. Cells **236**, 111519 (2022)

S8 Bach, G., Daehne, S.: Cyanine Dyes and Related Compounds

S9 Wöhrle, D., Schnurpfeil, G., Makarov, S.G., Kazarin, A., Suvorova, O.N.: Practical applications of phthalocyanines–from dyes and pigments to materials for optical, electronic and photo-electronic devices. Macroheterocycles **5**(3), 191–202 (2012)

S10 Liang, Y., Zhang, D., Wang, J., Lv, M., Yan, F., Xu, Z., Du, J., Jiang, X.-D.: A near-infrared absorbing ring-fused quinoid-bisboron pyrrole dye with double BF2 chelation. Dyes and Pigments **206**, 110613 (2022)

S11 de Campos Ventura-Camargo, B., Marin-Morales, M.A.: Azo dyes: characterization and toxicity-a review. Textiles and Light Industrial Science and Technology **2**(2), 85–103 (2013)

S12 Radulescu-Grad, M., Muntean, S., Todea, A., Verdes, O., Andelescu, A.: Synthesis and characterization of new metal complex dye. Chemical Bulletin of Politehnica University of Timisoara **60**(74), 37–40 (2015)

S13 Pérez-Juste, J., Pastoriza-Santos, I., Liz-Marzán, L.M., Mulvaney, P.: Gold nanorods: synthesis, characterization and applications. Coordin. Chem. Rev. **249**(17-18), 1870–1901 (2005)

S14 Marzouk, M., ElBatal, H., Hamdy, Y., Ezz-Eldin, F.: Collective optical, FTIR, and photoluminescence spectra of CeO_2 and/or Sm_2O_3-doped Na_2O–ZnO–P_2O_5 glasses. Int. J. Opt. **2019** (2019)

S15 Wang, F., Liao, Q., Chen, K., Pan, S., Lu, M.: Glass formation and FTIR spectra of CeO_2-doped $36\,Fe_2O_3$–$10\,B_2O_3$–$54\,P_2O_5$ glasses. J. Non-Cryst. Solids **409**, 76–82 (2015)

S16 Srinivasan, R., Yogamalar, R., Bose, A.C.: Structural and optical studies of yttrium oxide nanoparticles synthesized by co-precipitation method. Mater. Res. Bull. **45**(9), 1165–1170 (2010)

S17 Jacquemin, D., Perpete, E.A., Scuseria, G.E., Ciofini, I., Adamo, C.: TD-DFT performance for the visible absorption spectra of organic dyes: conventional versus long-range hybrids. J. Chem. Theory Comput. **4**(1), 123–135 (2008)

S18 Laurent, A.D., Adamo, C., Jacquemin, D.: Dye chemistry with time-dependent density functional theory. Phys. Chem. Chem. Phys. **16**(28), 14334–14356 (2014)

S19 Li, B., Zhao, M., Zhang, F.: Rational design of near-infrared-II organic molecular dyes for bioimaging and biosensing. ACS Mater. **2**(8), 905–917 (2020)

S20 Lewera, A., Timperman, L., Roguska, A., Alonso-Vante, N.: Metal–support interactions between nanosized Pt and metal oxides (WO3 and TiO2) studied using X-ray photoelectron spectroscopy. J. Phys. Chem. C **115**(41), 20153–20159 (2011)

S21 Arad, B., Ben-Shahar, O.: Filter selection for hyperspectral estimation. In: Proceedings of the IEEE International Conference on Computer

Vision, pp. 3153–3161 (2017)

S22 Labrín, C., Urdinez, F.: Principal component analysis, 375–393 (2020)

S23 Yuan, X., Brady, D.J., Katsaggelos, A.K.: Snapshot compressive imaging: Theory, algorithms, and applications. IEEE Signal Processing Magazine **38**(2), 65–88 (2021)

S24 Jalali, S., Yuan, X.: Snapshot compressed sensing: Performance bounds and algorithms. IEEE Transactions on Information Theory **65**(12), 8005–8024 (2019)

S25 Lammertyn, J., Peirs, A., De Baerdemaeker, J., Nicolaı, B.: Light penetration properties of NIR radiation in fruit with respect to non-destructive quality assessment. Postharvest Biol. Technol. **18**(2), 121–132 (2000)

S26 Li, J., Huang, W., Zhao, C., Zhang, B.: A comparative study for the quantitative determination of soluble solids content, pH and firmness of pears by Vis/NIR spectroscopy. J. Food Eng. **116**(2), 324–332 (2013)

S27 Leiva-Valenzuela, G.A., Lu, R., Aguilera, J.M.: Prediction of firmness and soluble solids content of blueberries using hyperspectral reflectance imaging. J. Food Eng. **115**(1), 91–98 (2013)

S28 Gitelson, A.A., Merzlyak, M.N., Lichtenthaler, H.K.: Detection of red edge position and chlorophyll content by reflectance measurements near 700 nm. J. Plant Physiol. **148**(3-4), 501–508 (1996)

S29 Shao, D., Liu, C., Tsow, F., Yang, Y., Du, Z., Iriya, R., Yu, H., Tao, N.: Noncontact monitoring of blood oxygen saturation using camera and dual-wavelength imaging system. IEEE.Trans. Biomed. Eng. **63**(6), 1091–1098 (2015)

S30 Ouyang, R., Wang, D., Jin, L., Zhang, X.: Dual-gratings imaging spectrometer. Appl. Sci. **11**(22), 11048 (2021)

S31 Stark, F., Rosenberger, M., Dittrich, P.-G., Celestre, R., Haensel, M., Notni, G.: Geometrical and spectral data stitching for combining hyperspectral imaging systems. At-autom. **65**(6), 381–395 (2017)

S32 Dixit, Y., Al-Sarayreh, M., Craigie, C., Reis, M.M.: A rapid method of hypercube stitching for snapshot multi-camera system. In: 2020 35th International Conference on Image and Vision Computing New Zealand

(IVCNZ), pp. 1–6 (2020)

S33 Zhou, J., Yu, L., Ding, Q., Wang, R.: Textile fiber identification using near-infrared spectroscopy and pattern recognition. Autex Res. J. **19**(2), 201–209 (2019)

S34 Liu, Z., Li, W., Wei, Z.: Qualitative classification of waste textiles based on near infrared spectroscopy and the convolutional network. Text. Res. J. **90**(9-10), 1057–1066 (2020)

S35 Le Moigne, J., Netanyahu, N.S., Eastman, R.D.: Image registration for remote sensing (2011)

S36 Lowe, D.G.: Object recognition from local scale-invariant features. In: Proceedings of the Seventh IEEE International Conference on Computer Vision, vol. 2, pp. 1150–1157 (1999)

S37 Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. **60**, 91–110 (2004)

S38 Products - Specim. https://www.specim.com/products/

S39 Products - Headwall Photonics. https://headwallphotonics.com/products/

S40 Hyperspectral Camera_FigSpec. https://www.figspec.com/en/h-col-141.html

S41 www.thorlabs.com. https://www.thorlabs.com/thorproduct.cfm?partnumber=CCS200#ad-image-0

S42 NIRQuest Spectrometer Series / Near-Infrared / Ocean Insight. https://www.oceaninsight.com/products/spectrometers/near-infrared/preconfigured-nirquest-spectrometers/

S43 Arad, B., Ben-Shahar, O.: Sparse recovery of hyperspectral signal from natural RGB images. In: Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part VII 14, pp. 19–34 (2016)

S44 Chakrabarti, A., Zickler, T.: Statistics of real-world hyperspectral images. In: CVPR 2011, pp. 193–200 (2011)

S45 Yasuma, F., Mitsunaga, T., Iso, D., Nayar, S.K.: Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum. IEEE Trans. Image Process. **19**(9), 2241–2253 (2010)

S46 Choi, I., Jeon, D.S., Nam, G., Gutierrez, D., Kim, M.H.: High-quality hyperspectral reconstruction using a spectral prior. ACM Trans. Graph. **36**(6), 1–13 (2017)

S47 Foster, D.H., Amano, K., Nascimento, S.M., Foster, M.J.: Frequency of metamerism in natural scenes. J. Opt. Soc. Am. A **23**(10), 2359–2372 (2006)

S48 Monno, Y., Teranaka, H., Yoshizaki, K., Tanaka, M., Okutomi, M.: Single-sensor RGB-NIR imaging: High-quality system design and prototype implementation. IEEE Sens. J. **19**(2), 497–507 (2018)

S49 Monno, Y., Kikuchi, S., Tanaka, M., Okutomi, M.: A practical one-shot multispectral imaging system using a single image sensor. IEEE Trans. Image Process. **24**(10), 3048–3059 (2015)

S50 Arad, B., Timofte, R., Yahel, R., Morag, N., Bernat, A., Cai, Y., Lin, J., Lin, Z., Wang, H., Zhang, Y., *et al.*: Ntire 2022 spectral recovery challenge and data set. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 863–881 (2022)

S51 Zhang, W., Song, H., He, X., Huang, L., Zhang, X., Zheng, J., Shen, W., Hao, X., Liu, X.: Deeply learned broadband encoding stochastic hyperspectral imaging. Light: Sci. Appl. **10**(1), 108 (2021)

S52 Cai, Y., Lin, J., Hu, X., Wang, H., Yuan, X., Zhang, Y., Timofte, R., Van Gool, L.: Mask-guided spectral-wise transformer for efficient hyperspectral image reconstruction. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 17502–17511 (2022)

S53 Cai, Y., Lin, J., Wang, H., Yuan, X., Ding, H., Zhang, Y., Timofte, R., Gool, L.V.: Degradation-aware unfolding half-shuffle transformer for spectral compressive imaging. Adv. Neur. In. **35**, 37749–37761 (2022)

S54 Zhang, W., Suo, J., Dong, K., Li, L., Yuan, X., Pei, C., Dai, Q.: Handheld snapshot multi-spectral camera at tens-of-megapixel resolution. Nat. Commun. **14**(1), 5043 (2023)

S55 Jeon, D.S., Baek, S.-H., Yi, S., Fu, Q., Dun, X., Heidrich, W., Kim, M.H.: Compact snapshot hyperspectral imaging with diffracted rotation (2019)

S56 Yako, M., Yamaoka, Y., Kiyohara, T., Hosokawa, C., Noda, A., Tack, K., Spooren, N., Hirasawa, T., Ishikawa, A.: Video-rate hyperspectral camera based on a CMOS-compatible random array of Fabry–Pérot filters. Nat. Photonics **17**(3), 218–223 (2023)

S57 Fu, Y., Zhang, T., Wang, L., Huang, H.: Coded hyperspectral image reconstruction using deep external and internal learning. IEEE Trans. Pattern Anal. Mach. Intell. **44**(7), 3404–3420 (2021)

S58 Zhang, S., Wang, L., Zhang, L., Huang, H.: Learning tensor low-rank prior for hyperspectral image reconstruction. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 12006–12015 (2021)

S59 Lawrence, K.C., Park, B., Windham, W.R., Mao, C.: Calibration of a pushbroom hyperspectral imaging system for agricultural inspection. T. ASAE **46**(2), 513 (2003)

S60 Amigo, J.M., Babamoradi, H., Elcoroaristizabal, S.: Hyperspectral image analysis. A tutorial. Anal. Chim. Acta. **896**, 34–51 (2015)

S61 Hruska, R., Mitchell, J., Anderson, M., Glenn, N.F.: Radiometric and geometric analysis of hyperspectral imagery acquired from an unmanned aerial vehicle. Remote Sens. **4**(9), 2736–2752 (2012)

S62 Wang, Z., Cun, X., Bao, J., Zhou, W., Liu, J., Li, H.: Uformer: A general u-shaped transformer for image restoration. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 17683–17693 (2022)

S63 Zamir, S.W., Arora, A., Khan, S., Hayat, M., Khan, F.S., Yang, M.-H.: Restormer: Efficient transformer for high-resolution image restoration. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 5728–5739 (2022)

S64 Hendrycks, D., Gimpel, K.: Gaussian error linear units (gelus). arXiv preprint arXiv:1606.08415 (2016)

S65 Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)

S66 Eason, D.T., Andrews, M.: Total variation regularization via continuation to recover compressed hyperspectral images. IEEE Trans. Image Process. **24**(1), 284–293 (2014)

S67 Wang, L., Xiong, Z., Huang, H., Shi, G., Wu, F., Zeng, W.: High-speed hyperspectral video acquisition by combining nyquist and compressive sampling. IEEE Trans. Pattern Anal. Mach. Intell. **41**(4), 857–870 (2018)

S68 Wang, L., Xiong, Z., Shi, G., Wu, F., Zeng, W.: Adaptive nonlocal sparse representation for dual-camera compressive hyperspectral imaging. IEEE Trans. Pattern Anal. Mach. Intell. **39**(10), 2104–2111 (2016)

S69 He, W., Yao, Q., Li, C., Yokoya, N., Zhao, Q., Zhang, H., Zhang, L.: Non-local meets global: An iterative paradigm for hyperspectral image restoration. IEEE Trans. Pattern Anal. Mach. Intell. **44**(4), 2089–2107 (2020)

S70 Liu, Y., Yuan, X., Suo, J., Brady, D.J., Dai, Q.: Rank minimization for snapshot compressive imaging. IEEE Trans. Pattern Anal. Mach. Intell. **41**(12), 2990–3006 (2018)

S71 Yuan, X.: Generalized alternating projection based total variation minimization for compressive sensing. In: 2016 IEEE International Conference on Image Processing (ICIP), pp. 2539–2543 (2016)

S72 Zheng, S., Liu, Y., Meng, Z., Qiao, M., Tong, Z., Yang, X., Han, S., Yuan, X.: Deep plug-and-play priors for spectral snapshot compressive imaging. Photonics Res. **9**(2), 18–29 (2021)

S73 Baraniuk, R.G., Cevher, V., Duarte, M.F., Hegde, C.: Model-based compressive sensing. IEEE Trans. Inf. Theory **56**(4), 1982–2001 (2010)

S74 Yang, Y., Sun, J., Li, H., Xu, Z.: ADMM-CSNet: A deep learning approach for image compressive sensing. IEEE Trans. Pattern Anal. Mach. Intell. **42**(3), 521–538 (2018)

S75 Shi, Z., Chen, C., Xiong, Z., Liu, D., Wu, F.: Hscnn+: Advanced cnn-based hyperspectral recovery from rgb images. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 939–947 (2018)

S76 Zhao, Y., Po, L.-M., Yan, Q., Liu, W., Lin, T.: Hierarchical regression network for spectral reconstruction from RGB images. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pp. 422–423 (2020)

S77 Cai, Y., Lin, J., Lin, Z., Wang, H., Zhang, Y., Pfister, H., Timofte, R., Van Gool, L.: Mst++: Multi-stage spectral-wise transformer for efficient spectral reconstruction. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 745–755 (2022)

S78 Hu, X., Cai, Y., Lin, J., Wang, H., Yuan, X., Zhang, Y., Timofte, R., Van Gool, L.: Hdnet: High-resolution dual-domain learning for spectral compressive imaging. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 17542–17551 (2022)

S79 Cai, Y., Lin, J., Hu, X., Wang, H., Yuan, X., Zhang, Y., Timofte, R., Van Gool, L.: Coarse-to-fine sparse transformer for hyperspectral image reconstruction. In: European Conference on Computer Vision, pp. 686–704 (2022)

S80 Yesilkoy, F., Arvelo, E.R., Jahani, Y., Liu, M., Tittl, A., Cevher, V., Kivshar, Y., Altug, H.: Ultrasensitive hyperspectral imaging and biodetection enabled by dielectric metasurfaces. Nat. Photonics **13**(6), 390–396 (2019)

S81 Association, E.M.V., et al.: Standard for characterization of image sensors and cameras. EMVA Standard **1288** (2010)

S82 Hakkel, K.D., Petruzzella, M., Ou, F., van Klinken, A., Pagliano, F., Liu, T., van Veldhoven, R.P., Fiore, A.: Integrated near-infrared spectral sensing. Nat. Commun. **13**(1), 103 (2022)

S83 Skauli, T.: Characterizing the dynamic range of a hyperspectral camera. In: Algorithms, Technologies, and Applications for Multispectral and Hyperspectral Imaging XXVII, vol. 11727, pp. 40–45 (2021)

S84 Bao, J., Bawendi, M.G.: A colloidal quantum dot spectrometer. Nature **523**(7558), 67–70 (2015)

S85 Plaza, A., Benediktsson, J.A., Boardman, J.W., Brazile, J., Bruzzone, L., Camps-Valls, G., Chanussot, J., Fauvel, M., Gamba, P., Gualtieri, A., *et al.*: Recent advances in techniques for hyperspectral image processing. Remote Sens. Environ. **113**, 110–122 (2009)

S86 CMS series Multispectral Cameras (silios.com). https://www.silios.com/cms-series

S87 Wagadarikar, A., John, R., Willett, R., Brady, D.: Single disperser design for coded aperture snapshot spectral imaging. Appl. Opt. **47**(10), 44–51 (2008)

S88 Arguello, H., Arce, G.R.: Colored coded aperture design by concentration of measure in compressive spectral imaging. IEEE T. Image Process. **23**(4), 1896–1908 (2014)

S89 Wagadarikar, A.A., Pitsianis, N.P., Sun, X., Brady, D.J.: Video rate spectral imaging using a coded aperture snapshot spectral imager. Optics express **17**(8), 6368–6388 (2009)

S90 Lin, X., Liu, Y., Wu, J., Dai, Q.: Spatial-spectral encoded compressive hyperspectral imaging. ACM T. Graph. **33**(6), 1–11 (2014)

S91 Yang, J., Cui, K., Cai, X., Xiong, J., Zhu, H., Rao, S., Xu, S., Huang, Y., Liu, F., Feng, X., *et al.*: Ultraspectral imaging based on metasurfaces with freeform shaped meta-atoms. Laser Photonics Rev. **16**(7), 2100663 (2022)

S92 Lalanne, P., Hugonin, J.P., Chavel, P.: Optical properties of deep lamellar gratings: a coupled Bloch-mode insight. J. Light. Technol. **24**(6), 2442 (2006)

S93 Karagodsky, V., Chang-Hasnain, C.J.: Physics of near-wavelength high contrast gratings. Opt. Express **20**(10), 10888–10895 (2012)

S94 Sedgwick, P.: Pearson's correlation coefficient. Brit. Med. J. **345** (2012)

S95 Abgrall, P., Conedera, V., Camon, H., Gue, A.-M., Nguyen, N.-T.: SU-8 as a structural material for labs-on-chips and microelectromechanical systems. Electrophoresis **28**(24), 4539–4551 (2007)

S96 Martinez-Duarte, R., Madou, M.: SU-8 photolithography and its impact on microfluidics. Microfluid. Nanofluidics (2006), 231–268 (2011)

S97 MicroChem: SU-8 2000 Permanent Epoxy Negative Photo Resist. MicroChem, Inc. Newton, MA (2007)

S98 Youn, S.-W., Ueno, A., Takahashi, M., Maeda, R.: Microstructuring of SU-8 photoresist by UV-assisted thermal imprinting with non-transparent mold. Microelectron. Eng. **85**(9), 1924–1931 (2008)

S99 Hikrobot: MV-CA050-12UM/UC Technical Manual. https://www.maxxvision.com/downloads/Cameras/USB30/Hikrobot/Hikrobot_MV-CA050-12UMUC.pdf (2023)

S100 Evennett, P.: Kohler illumination: a simple interpretation. R. Microsc.

Soc. Proc. **28**(4), 189–192 (1983)

S101 Williams, P., Norris, K., et al.: Near-infrared technology in the agricultural and food industries. (1987)

S102 Jamshidi, B., Minaei, S., Mohajerani, E., Ghassemian, H.: Reflectance Vis/NIR spectroscopy for nondestructive taste characterization of Valencia oranges. Comput. Electron. Agric. **85**, 64–69 (2012)

S103 Likas, A., Vlassis, N., Verbeek, J.J.: The global k-means clustering algorithm. Pattern Recognit. **36**(2), 451–461 (2003)

S104 Li, W., Wei, Z., Liu, Z., Du, Y., Zheng, J., Wang, H., Zhang, S.: Qualitative identification of waste textiles based on near-infrared spectroscopy and the back propagation artificial neural network. Text. Res. J. **91**(21-22), 2459–2467 (2021)

S105 Kennard, R.W., Stone, L.A.: Computer aided design of experiments. Technometrics, 137–148 (1969)

S106 Cortes, C., Vapnik, V.: Support-vector networks. Mach. Learn. **20**, 273–297 (1995)

S107 Riba, J.-R., Cantero, R., Riba-Mosoll, P., Puig, R.: Post-consumer textile waste classification through near-infrared spectroscopy, using an advanced deep learning approach. Polymers **14**(12), 2475 (2022)

S108 Yogesh, Dubey, A.K., Ratan, R., Rocha, A.: Computer vision based analysis and detection of defects in fruits causes due to nutrients deficiency. Cluster Comput. **23**, 1817–1826 (2020)

S109 Hussein, Z., Fawole, O.A., Opara, U.L.: Harvest and postharvest factors affecting bruise damage of fresh fruits. Hortic. Plant J. **6**(1), 1–13 (2020)

S110 Tian, X., Liu, X., He, X., Zhang, C., Li, J., Huang, W.: Detection of early bruises on apples using hyperspectral reflectance imaging coupled with optimal wavelengths selection and improved watershed segmentation algorithm. J. Sci. Food Agric. **103**(13), 6689–6705 (2023)

S111 Pang, Q., Huang, W., Fan, S., Zhou, Q., Wang, Z., Tian, X.: Detection of early bruises on apples using hyperspectral imaging combining with YOLOv3 deep learning algorithm. J. Food Process Eng. **45**(2), 13952 (2022)

S112 Jocher, G., Stoken, A., Chaurasia, A., Borovec, J., Kwon, Y., Michael, K., Changyu, L., Fang, J., Skalski, P., Hogan, A., et al.: ultralytics/yolov5: v6. 0-yolov5n'nano'models, roboflow integration, tensorflow export, opencv dnn support. Zenodo (2021)