

SpectraTrack: megapixel, hundred-fps, and thousand-channel hyperspectral imaging

Received: 27 March 2024

Accepted: 17 October 2024

Published online: 01 November 2024

 Check for updatesDaoyu Li , Jinxuan Wu, Jiajun Zhao, Hanwen Xu & Liheng Bian  


Hyperspectral imaging (HSI) finds broad applications in various fields due to its substantial optical signatures for the intrinsic identification of physical and chemical characteristics. However, it faces the inherent challenge of balancing spatial, temporal, and spectral resolution due to limited bandwidth. Here we present *SpectraTrack*, a computational HSI scheme that simultaneously achieves high spatial, temporal, and spectral resolution in the visible-to-near-infrared (VIS-NIR) spectral range. We deeply investigated the spatio-temporal-spectral multiplexing principle inherent in HSI videos. Based on this theoretical foundation, the *SpectraTrack* system uses two cameras including a line-scan imaging spectrometer for temporal-multiplexed hyperspectral data and an auxiliary RGB camera to capture motion flow. The motion flow guides hyperspectral reconstruction by reintegrating the scanned spectra into a 4D video. The *SpectraTrack* system can achieve around megapixel HSI at 100 fps with 1200 spectral channels, demonstrating its great application potential from drone-based anti-vibration video-rate HSI to high-throughput non-cooperative anti-spoofing.

Hyperspectral imaging (HSI)^{1–3} acquires fine-grained spectrum of each spatial location, producing a hyperspectral datacube that contains abundant spectral characteristics of target scenes. The additional spectral information helps distinguish different material compositions of the same color to differentiate their physical and chemical properties. Benefiting from its capabilities beyond the human vision, HSI finds broad applications in diverse fields such as biomedical engineering^{4–7}, medical diagnosis^{8,9}, ecology¹⁰, agricultural inspection^{11,12}, machine vision^{13,14}, etc.

Early HSI techniques capture 3D hyperspectral datacubes under the spatial^{15,16} or spectral^{17,18} scanning architecture. While they can provide high spatial-spectral resolution in exchange with temporal resolution, they are not suitable for imaging dynamic scenes. Fang et al.¹⁹ employed an acousto-optic tunable filter and a high-speed image sensor to achieve video-rate spectral scanning HSI. To enable video-rate HSI, snapshot HSI techniques were developed to map the 3D spatial-spectral datacube into a 2D measurement by direct or compressed acquisition methods. The direct acquisition methods, such as prism-mask-based²⁰, lenslet-array-based²¹, spatial-light-

modulator-based²², metasurface-based²³, hyperspectral-filter-array-based²⁴ ones, etc., establish a one-to-one correspondence between the 3D datacube and 2D camera pixels. Nevertheless, they face an inherent limitation that the amount of information obtained in a single measurement does not exceed the number of pixels in the sensor. Therefore, all of these techniques suffer from an intrinsic tradeoff between spatial and spectral resolution, which impedes their practical applications. On the other hand, compressed acquisition techniques, including computed-tomography-based methods^{25,26}, amplitude^{27–31}, phase^{32,33}, and wavelength-coded^{34,35} approaches, compress hyperspectral datacubes along the spectral dimension into 2D measurements. Then hyperspectral images can be subsequently recovered from these under-sampled measurements using optimization-based³⁶ or learning-based^{36,37} reconstruction methods. Inherent to the compressive sensing theory, most of the existing compressed-acquisition-based prototypes reconstruct hyperspectral images with up to tens of spectral channels³⁸. Compressing more spectral channels within one measurement results in lower sampling ratios, making it more challenging for high-resolution reconstruction.

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

 e-mail: bian@bit.edu.cn

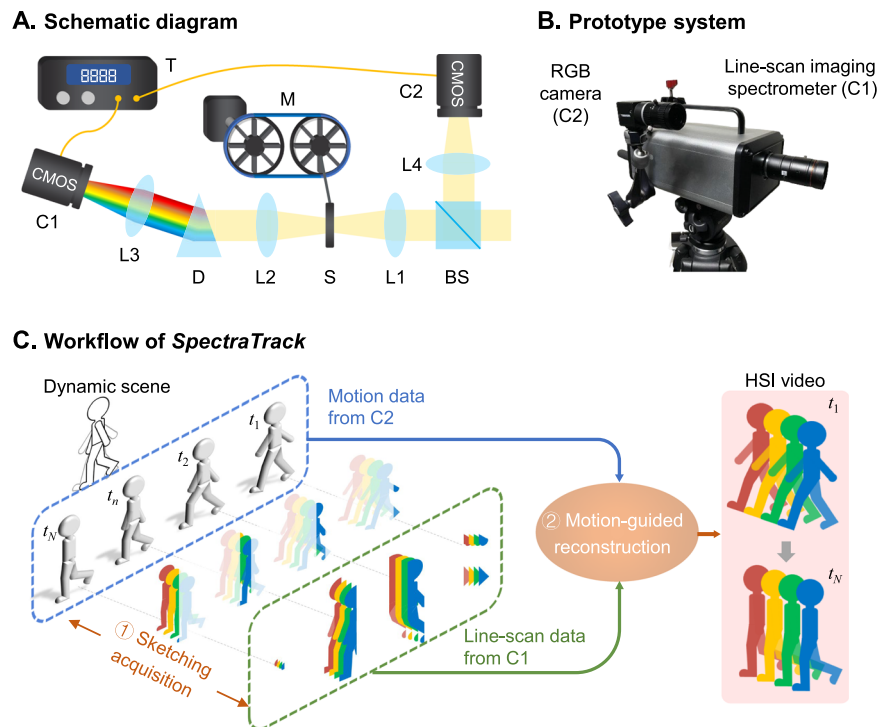


Fig. 1 | Principle of SpectraTrack. **A** The schematic diagram of the reported framework. The imaging optical path includes two branches, one for the line-scan imaging spectrometer, and the other for the array camera. C1, C2: CMOS sensor. L1–L4: lens. D: prism. S: slit. M: motor. T: trigger source. BS: beam splitter, indicated as the alignment of FOVs of C1 and C2. The FOVs of C1 and C2 can be aligned by optical setup or algorithmic calibration. **B** The prototype system is built by a side-

by-side placed line-scan imaging spectrometer and RGB camera. **C** The workflow of SpectraTrack, including sketching acquisition and motion-guided reconstruction. 1) Sketching acquisition: capturing the multiplexed HSI measurement by the imaging spectrometer and motion information by the RGB camera. 2) Motion-guided reconstruction: Reconstructing the 4D HSI video from HSI measurement with the guidance of motion information from the RGB camera video.

In this work, we report a computational HSI scheme of high spatio-temporal-spectral resolution, termed *SpectraTrack*. This technique investigates the strong spatio-temporal-spectral redundancy in hyperspectral videos, and derives a computational imaging modality following multidimensional sketching acquisition and motion-guided hyperspectral restoration. *SpectraTrack* integrates a line-scan imaging spectrometer and an RGB camera. The spectrometer records spatio-temporal-coupled 3D line-scan data, while the RGB camera tracks target motion, enabling the reconstruction of 4D hyperspectral videos by matching captured spectra to their spatial positions at each acquisition time. *SpectraTrack* significantly enhances frame rates by more than two orders of magnitude for scanning-based imaging spectrometers, aligning with the speed of high-speed RGB cameras. It can achieve approximately megapixel HSI at 100 fps with 1200 spectral channels spanning 400–1000 nm. We experimentally demonstrated its applications from drone-based video-rate HSI to high-throughput anti-spoofing, where *SpectraTrack* excels in precision and spectral range compared to conventional systems. More comparisons between *SpectraTrack* and existing HSI techniques can be found in Supplementary Note 1. Overall, this technique breaks the tradeoff among spatial, temporal, and spectral resolution. We anticipate that this technique expands the possibilities for high-throughput HSI and facilitates broad application scenarios.

Results

SpectraTrack

Spatial scanning imaging spectrometers are typically limited in their imaging speed due to the use of mechanical scanning, allowing for HSI mostly in stationary scenes. If there is a relative displacement between the spatial scanning imaging spectrometer and the object in the imaging process, adverse effects such as artifacts and stretching will

appear in the acquired images. We found that these effects arise from the blending of scene information from multiple moments within one measurement (Fig. 1C). Each spectrum in the scanned HSI measurement corresponds to a specific acquisition moment. We can reconstruct a complete HSI video of all acquisition moments by establishing the inverse process of the acquisition. Assuming that no variation occurs in the spectral properties of the scene materials or the illumination during the imaging process, this inverse process is to find the spatial positions corresponding to all spectra in the scanned HSI measurement at all moments. Therefore, the reconstruction requires the incorporation of scene motion information, which can be achieved by introducing a common video-rate RGB camera aligned with the scanning-based imaging spectrometer (Fig. 1A, B). Then we can reconstruct the 4D HSI video using the scanned HSI measurement and the dense motion flow (Figs. 1C and 2B). The reported *SpectraTrack* technique differs from the existing spatial scanning¹⁶, spectral scanning³⁹, and coded aperture²⁷ systems (Fig. 2A) in that it leverages the inherent information multiplexing in the spatio-temporal dimensions. Details can be found in the Method section.

System architecture

As shown in Fig. 1A, B, the prototype includes a high-speed RGB camera (Hikrobot MV-CA013-AOUC: 1280 × 1024 pixels, up to 200 fps) and a line-scan imaging spectrometer (FigSpec Hyperspectral Camera FS-23: up to 1920 pixels per line, 400–1000 nm, 1200 channels, 2.5 nm spectral resolution) placed side by side. We matched their field of view (FOV) through prior calibration and post-registration refinement. Our prototype first synchronously triggers both the scanning imaging spectrometer and RGB camera using a triggering source to capture line-scan data and RGB videos, respectively. Next, with the guidance of motion information extracted from the RGB video, we reconstruct the

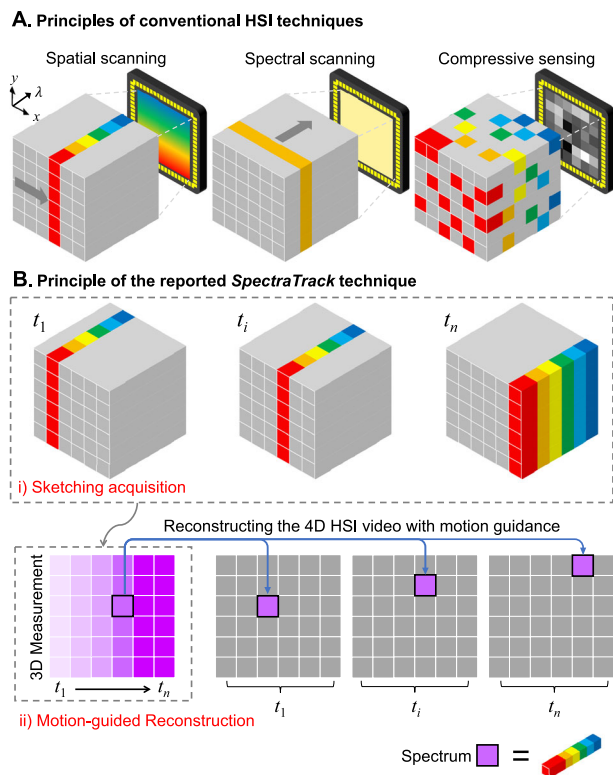


Fig. 2 | Comparison between conventional HSI techniques and the reported *SpectraTrack*. **A** Principles of the conventional HSI frameworks including spatial-scanning-based, spectral-scanning-based, and compressed-acquisition-based ones. **B** Principle of the reported *SpectraTrack* technique. We first capture the multiplexed HSI measurement of dynamic scenes via spatial scanning. Each row of the measurement corresponds to a slice of the target's hyperspectral information cube at a moment. The measurement was then utilized to reconstruct the spectra at all trigger timings. By finding the spatial positions of all scanned spectra at all trigger timings, we can reconstruct the hyperspectral image at each timing to form a 4D HSI video. The spatial, temporal, and spectral resolutions of the system are constrained by the imaging hardware. In the reconstructed 4D HSI video $D_{T,X,Y,C}$, T (time) and X (horizontal) are determined by the number of scanning iterations, while Y (vertical) and C (channel) are defined by the spatial resolution of the imaging sensor in the spectrometer. The optical configuration of the imaging spectrometer determines the spectral range and resolution. If the data transmission requirements are met, the system frame rate depends on the minimum frame rate of the imaging sensors in the imaging spectrometer and the RGB camera. Constrained by data transmission efficiency (see Supplementary Note 5), we currently achieve high-throughput HSI with a resolution of 960×960 pixels, a frame rate of 100 fps, 1200 spectral channels, and a spectral range of 400–1000 nm.

4D HSI video from the line-scan data, as depicted in Figs. 1C and 2. Figure 2B demonstrates that each spectrum in the line-scan data corresponds to a trigger timing. We can calculate the scene motion flow between this trigger timing and other trigger timings to predict the spatial position corresponding to this spectrum at other trigger timings. By finding the spatial positions of all scanned spectra at all trigger timings, we can reconstruct the hyperspectral image at each timing to form a 4D HSI video. The spatial, temporal, and spectral resolutions of the system are constrained by the imaging hardware. In the reconstructed 4D HSI video $D_{T,X,Y,C}$, T (time) and X (horizontal) are determined by the number of scanning iterations, while Y (vertical) and C (channel) are defined by the spatial resolution of the imaging sensor in the spectrometer. The optical configuration of the imaging spectrometer determines the spectral range and resolution. If the data transmission requirements are met, the system frame rate depends on the minimum frame rate of the imaging sensors in the imaging spectrometer and the RGB camera. Constrained by data transmission efficiency (see Supplementary Note 5), we currently achieve high-throughput HSI with a resolution of 960×960 pixels, a frame rate of 100 fps, 1200 spectral channels, and a spectral range of 400–1000 nm.

High-throughput HSI for dynamic scenes

To demonstrate the advanced dynamic high-resolution HSI capability of the system, we utilized the above settings and built a scene with multiple moving objects, as shown in Fig. 3. We would

up the windmill to rotate automatically, and manually pulled the toy train. The toy train moved from right to left, while the imaging spectrometer scanned from left to right. So we can observe the shrinking effect in the horizontal direction of the toy train in measurement (Fig. 3A). The windmill rotated counter-clockwise, which led to distortion for conventional hyperspectral imagery (Fig. 3A). The reconstructed HSI video of the *SpectraTrack* system is shown in Fig. 3D, Supplementary Video 1, and Supplementary Note 8, which show clear motion of the moving and rotating targets.

HSI channels corresponding to (652.6, 525.7, 481.0) nm are applied for the pseudo-RGB images of reconstructions. In the reconstructed hyperspectral video, the boundaries of each grid on the blades and the textures on the train are clear and sharp. Our *SpectraTrack* technique can reconstruct the scene of the windmill rotating and the train moving in a manner consistent with those recorded in the RGB video. Moreover, the reported system has high hyperspectral accuracy, being capable of recording target spectra with 1200 VIS-NIR bands. It can accurately distinguish the difference between blue dyes on the train and the ColorChecker. The blue dye on the train has an obvious peak at around 800 nm, while the blue on the ColorChecker does not (Fig. 3C). If data acquisition is performed using high-speed acquisition cards, it can fully leverage the advantages of our optical setup, allowing for HSI with significantly surpassing 100 fps and 960×960 resolution. In summary, the system achieves high spatio-temporal-spectral resolution, which enables great application potential.

Drone-based anti-vibration video-rate HSI

Drone-based hyperspectral remote imaging⁴⁰ has become a vital tool in various fields, including precision agriculture, environmental monitoring, mineral exploration, urban planning, etc. During the flight, motors and airflow usually make vibrations, which can degrade the imaging quality. Therefore, drone-based measurement equipment is typically mounted on gimbals to reduce vibration. However, at higher altitudes where wind speeds are greater and airflow is less stable, even a gimbal may not completely counteract the effects of vibration, leading to significant impacts on high-precision spectral imaging equipment. As shown in Fig. 4A, with the wind speed of level 5 observed at the ground station, the drone (at a height of 100 m) experiences shake due to airflow, resulting in significant artifacts in the images captured by the line-scan HSI imager mounted with a gimbal. To address this challenge, the reported video-rate HSI system can be employed to remove the shaken-caused distortion and capture HSI videos. As shown in Fig. 4B, we placed the line-scan hyperspectral imager and RGB camera (both working at 35 fps) side by side and performed calibration and registration. The drone used in the experiment was the FlytoUAV Chata-5. The drone directly carried the reported video-rate HSI system without gimbal stabilization to an altitude of approximately 100 m, acquiring ground hyperspectral data from above.

Figure 4 demonstrates the vibration-affected drone-based HSI measurements and the reconstructed HSI videos. We tested the HSI video reconstruction under both aerial flight and hovering conditions, with results displayed as pseudo-RGB in Fig. 4C. The full videos are shown in Supplementary Video 2 and Supplementary Video 3. During the drone flight, vibrations caused distortions in the line-scan measurements. Our method can reconstruct clear spectral data of scenes in motion from distorted line-scan data. The measurements taken during drone hovering contained distortions in multiple directions, whereas our method can eliminate distortions in all directions and reconstruct the clear HSI video. By employing the reported method, the stability and speed of drone-based HSI can be significantly improved, demonstrating great potential for wide applications in drone-based remote imaging.

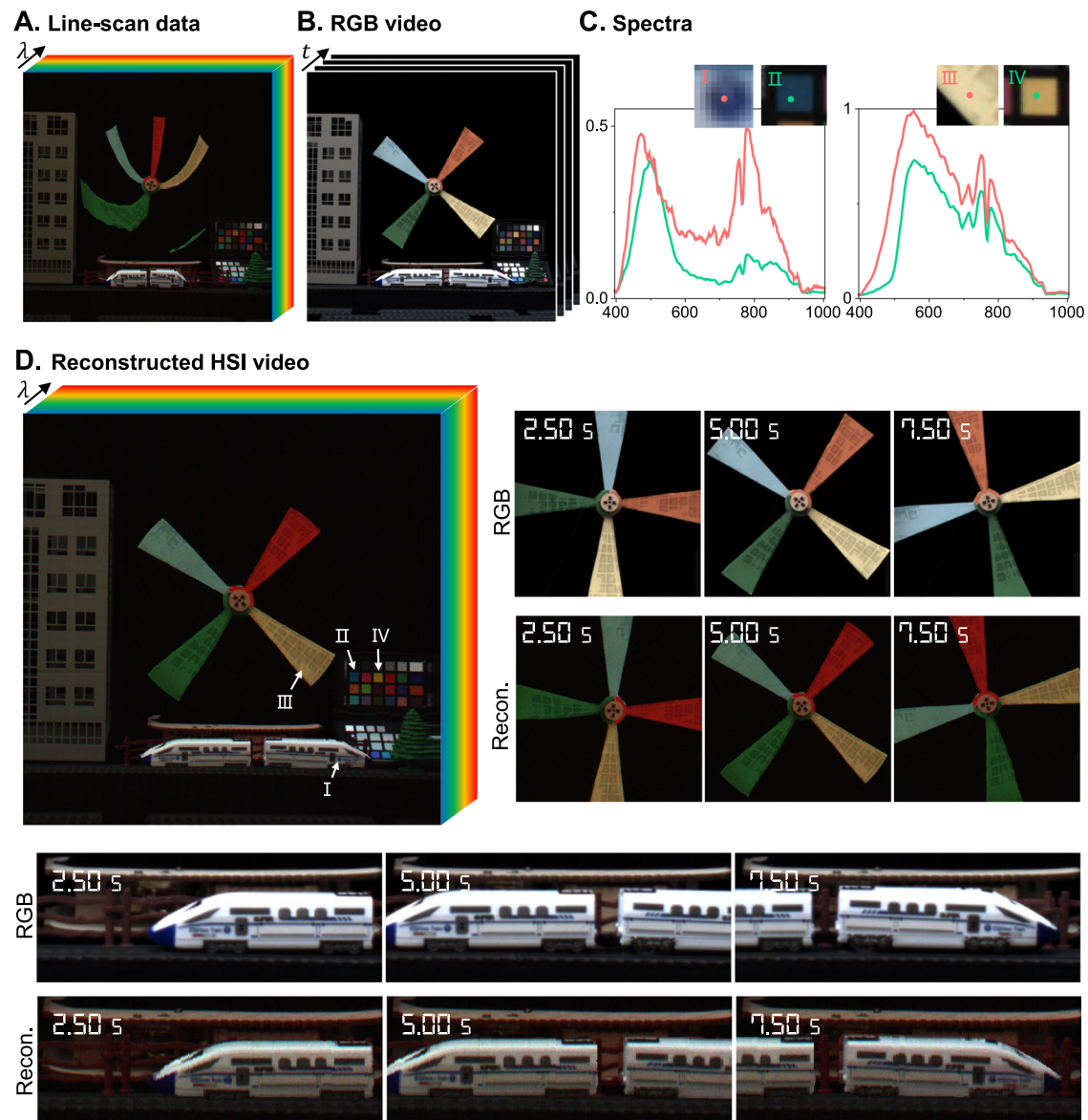


Fig. 3 | Imaging results of the dynamic scene. **A** Line-scan-data. **B** Captured RGB video. **C** Spectra of the objects and ColorChecker in **D**. **D** The reconstructed hyperspectral video (shown in pseudo-RGB format). Our prototype can reconstruct the clear motion of targets from the line-scan data.

High-throughput non-cooperative anti-spoofing

Anti-spoofing plays a vital role in safeguarding security in various situations, including airports, financial transactions, privacy protection, etc. Even human eyes can be deceived by intricate disguises featuring finely crafted shapes, colors, and overall appearance⁴¹. Whereas, these disguises usually inherently differ in their subtle spectral distribution due to different materials compared with real targets⁴². Therefore, HSI techniques can be used to identify disguises. In non-cooperative security inspection applications, distinguishing between real and fake skin of moving pedestrians requires not only high-precision spectral imaging but also dynamic imaging capability⁴². However, there is a fundamental trade-off between the spectral and temporal resolutions for existing HSI systems. In general, scanning-based HSI can provide higher spectral resolution than snapshot spectral imaging techniques. However, it typically requires more time to scan the scene, which may not meet the dynamic imaging requirement. During the scanning process, subjects' motion may cause artifacts, distortion, and other issues, degrading the imaging quality and detection accuracy. In comparison, our *SpectraTrack* method can

reconstruct high-precision spectral information of dynamic scenes, thus enabling non-cooperative HSI anti-spoofing, which is difficult to achieve using conventional HSI techniques.

To address this challenge, we employ the reported video-rate HSI system for high-throughput anti-spoofing. We first captured the spectra of the background. It helps recover the occluded portions caused by the pedestrian's presence when reconstructing the scenes of pedestrians' trajectories. The pedestrians walked in front of the system, with a few of them wearing disguised (silicone) masks. We have complied with all pertinent ethical regulations and secured informed consent from all four volunteers. Using the system, we aimed to distinguish the disguises. The imaging spectrometer scanned in the opposite direction from the walking direction of pedestrians. Figure 5A, B demonstrates the reconstructed HSI videos from corresponding line-scan data. Our system exhibited its capability of reconstructing clear HSI videos for multi-object motion scenarios. Furthermore, the spectra extracted from the reconstructed HSI videos provide a clear distinction between the spectral characteristics of real

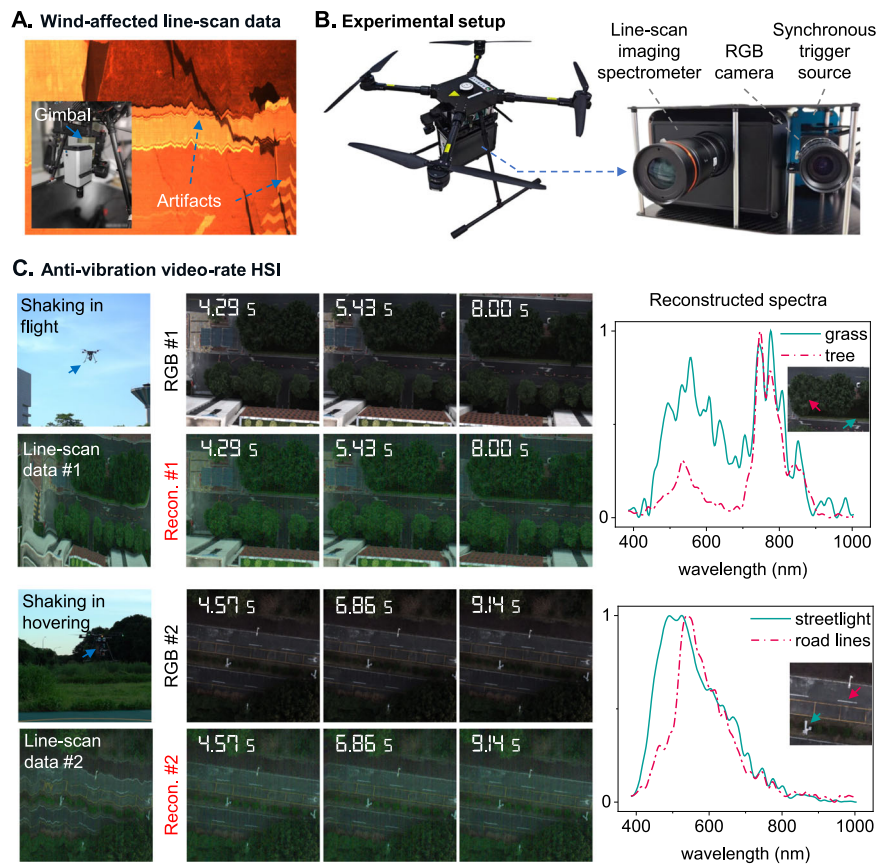


Fig. 4 | Drone-based anti-vibration video-rate HSI. **A** The imaging result of the line-scan imaging spectrometer installed on the drone with a gimbal. A gust up to level 5 causes shaking and leads to artifacts in the HSI image. **B** The reported system, comprising a side-by-side placed line-scan image spectrometer and an RGB

camera was mounted under the drone. **C** The shaking-affected HSI images and reconstructions in the cases of shaking in flight and hovering. Our system can reconstruct clear, high-resolution HSI videos and differentiate between the subtle spectra of objects that share similar colors.

and silicone skins. The spectra of silicone skin are significantly higher than those of real skin in the wavelength range of approximately 480–620 nm. While commercial snapshot multispectral cameras are capable of real-time imaging, their low spectral resolution makes it difficult to differentiate between real and disguised skins, as depicted in Fig. 5C. Experiments validated that the reported system simultaneously fulfills the spatial, temporal, and spectral resolution requirements for high-throughput anti-spoofing. This system exhibits the potential for further extension and application.

Applicability and limitation

The *SpectraTrack* technique is founded on three key prerequisites: (1) consistent illumination, (2) invariant spectral characteristics of the target, and (3) the motion and spectral data of the target are captured. In practice, these conditions are generally met. However, in a few instances, such as if conditions (1) and (2) are not met, when the imaging spectrometer scans each line, it can only record the illumination and spectral information at the moment of scanning, resulting in incorrect brightness or spectral information in the reconstructed hyperspectral images. Additionally, changes in illumination conditions or the target's color could potentially degrade the accuracy of extracted optical flows. Failure to meet condition (3), whether caused by occlusions or the target moving too fast, will lead to the line-scan imaging spectrometer missing the target's spectral information or the RGB camera being incapable of acquiring the target's motion information, thereby resulting in reconstruction failure. More details are in Supplementary Note 7.

The *SpectraTrack* technique works when the spectrometer captures targets' reflected light during scanning. On the other hand, the captured data will lose information if the target is moving too fast. Assume that the spectrometer scans v_s pixels per second and the object moves v_t pixels per second on the image plane. Decompose v_t into components v_h (parallel to v_s) and v_v (perpendicular to v_s). v_h is positive when aligned with v_s , and negative when opposite. Suppose that $r_h = v_h/v_s$ and $r_v = v_v/v_s$. The information retention ratio Ω in the captured line-scan data under various moving speeds and directions of the target is shown in Fig. 6A which indicates a conclusion consistent with intuition: as $|r|$ increases, Ω decreases. When $r_h > 1$, it indicates that the scanning speed cannot catch up with the target's motion speed; hence, $\Omega = 0$. When $|r_h|$ ($r_h < 0$) or $|r_v|$ increases, Ω tends towards 0, although theoretically, it would not be exactly 0, as at least one line would contain the target information. As shown in Fig. 6B, different motion directions will lead to different types of information loss. When v_s is opposite to v_t , the target in the line-scan data will be compressed, resulting in raster-like information loss in the reconstructed images. This can be later mitigated through inpainting. Unless otherwise specified, the results presented in this work have all utilized `cv2.inpaint` of OpenCV to enhance reconstruction quality. When v_s is parallel/perpendicular to v_t , the target in the line-scan data will be stretched/sheared, resulting in square-shaped/triangular-shaped information loss in the reconstructed images. The more complex motions can be considered as combinations of the aforementioned cases. More details are referred to Supplementary Note 6.

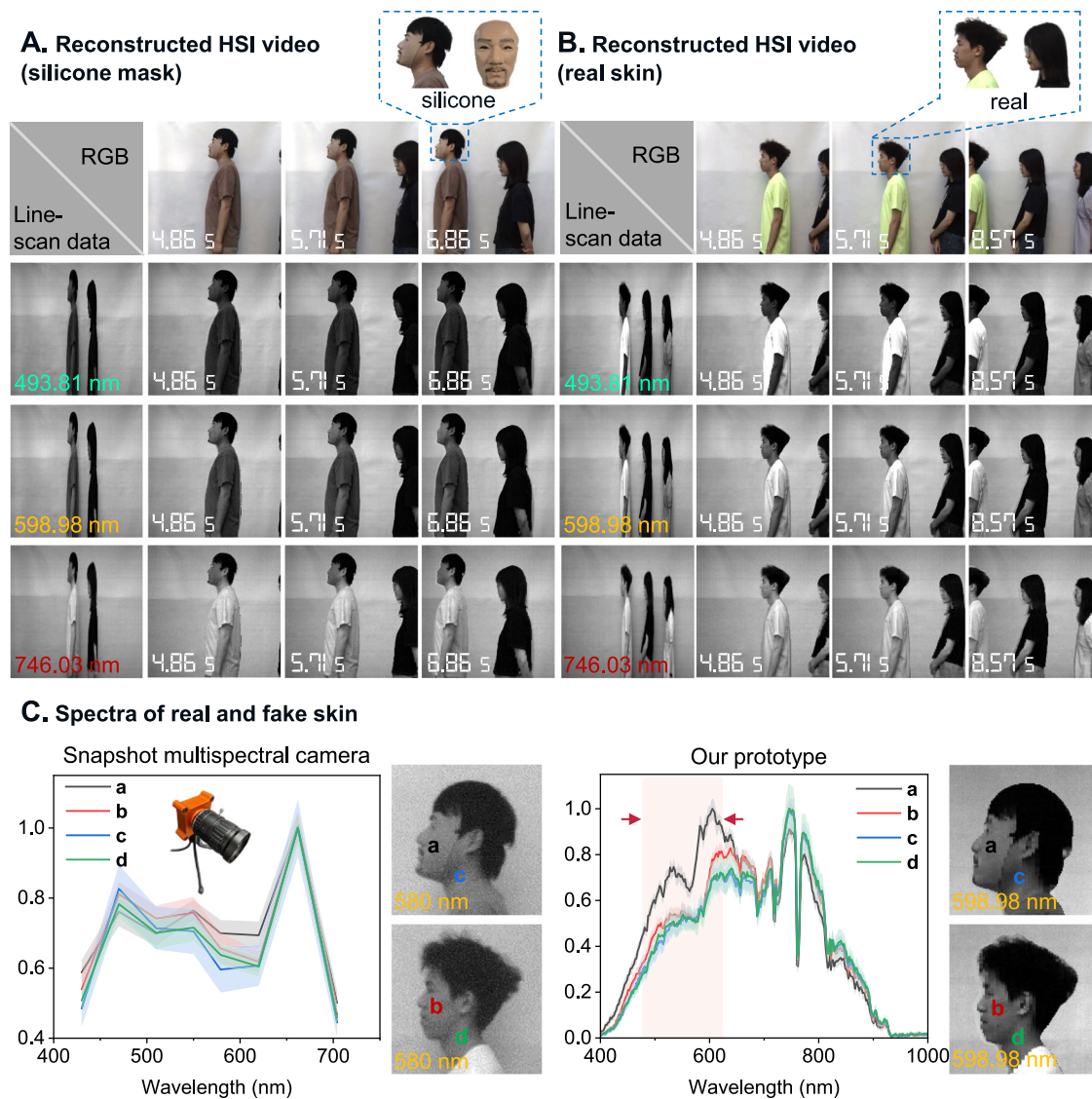


Fig. 5 | High-throughput non-cooperative anti-spoofing. **A** and **B** are the reconstructed HSI video with and without disguises. **C** The spectra of real and fake skins extracted from the videos obtained by our prototype and a commercial snapshot multispectral camera. (CMS-C, 3×3 channels). **a** Disguised skin made

from silicone. **b–d** Real skin. Our prototype can distinguish the difference in spectra between real and disguised skins. However, the snapshot multispectral camera cannot faithfully distinguish between the two due to lower spectral resolution.

Discussion

This study presents a HSI framework *SpectraTrack*, achieving around megapixel, hundred-fps, and thousand-channel HSI. By fully exploiting the advantages of integrating both a spatial-scanning HSI imager and a conventional RGB camera, *SpectraTrack* provides an innovative scheme for future ultra-high-precision video-rate HSI. Different from conventional hyperspectral imaging approaches that directly scan scenes, our method leverages the temporal-multiplexing characteristic of line-scan imaging spectrometers and reconstructs HSI videos through computational decoupling. We applied the *SpectraTrack* technique for pioneering tasks including drone-based anti-vibration video-rate HSI and high-throughput non-cooperative anti-spoofing, demonstrating its potential in extending the capabilities and applications of HSI.

In addition, we can promote performance and broaden applications of the reported framework through the following approaches.

Non-uniform scanning: In scenarios where the scene is sparse, utilizing non-uniform scanning can be highly effective. By employing

the RGB camera to detect sparse areas, the trigger speed can be reduced specifically in those regions. This approach is advantageous as it minimizes the collection of redundant spatial information, making the scanning process more efficient in terms of reducing data volumes. Compared to the uniform scanning method, non-uniform scanning is promising to be a more targeted and efficient approach.

Asynchronous acquisition: Our system currently synchronizes the line-scan imaging spectrometer and the RGB camera, where the former captures spectral information and the latter motion details. By transitioning to asynchronous acquisition, the RGB camera can adopt a higher frame rate, enabling the capture of more nuanced motion. If the RGB camera and the line-scan spectrometer begin or end simultaneously, it is possible to correlate RGB videos with the line-scan data for further HSI video reconstruction. This is particularly useful for overcoming temporal resolution limitations that are imposed by the hyperspectral scanning speed. Asynchronous acquisition enables a more comprehensive and detailed representation of motion in the reconstructed video.

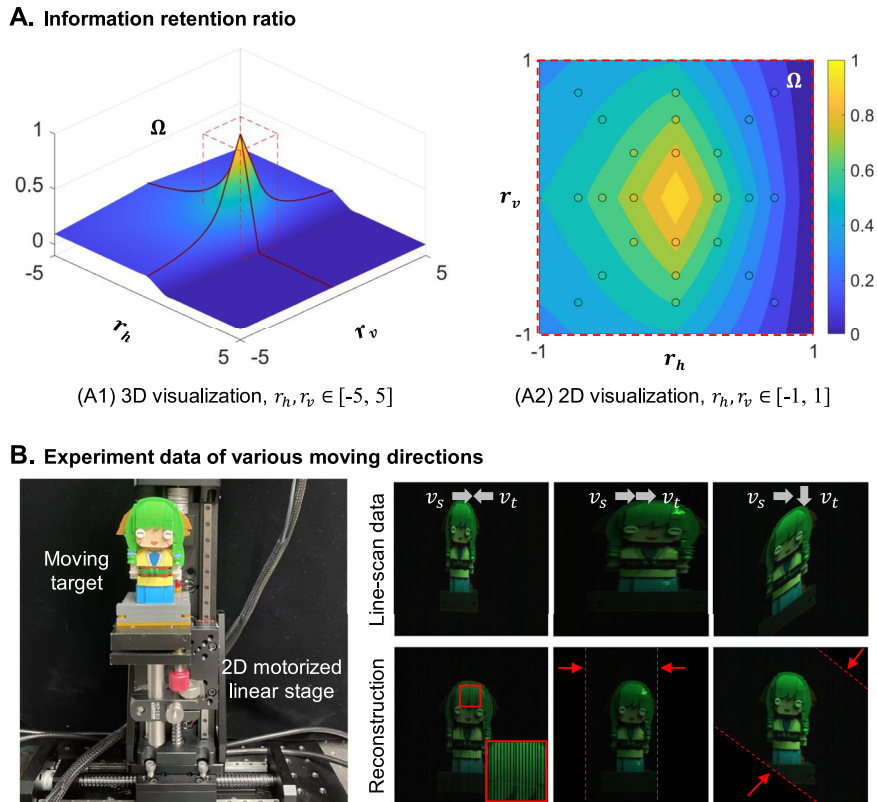


Fig. 6 | Information retention ratio and reconstructed results under various object moving directions and speeds. **A** The information retention ratio Ω w.r.t. r_h, r_v . (A1) 3D visualization, $r_h, r_v \in [-5, 5]$. The red curves indicate the theoretical results in the cases where v_t is in the same direction ($r_h > 0, r_v = 0$), opposite direction ($r_h < 0, r_v = 0$), or perpendicular ($r_h = 0$) to v_s . (A2) 2D visualization,

$r_h, r_v \in [-1, 1]$. The points represent the experimental results. **B** The experiment data of various moving directions. In the cases of relative, unidirectional, and perpendicular motions, the line-scan data and reconstructed image will exhibit different types of information loss. The region between the two red dashed lines indicated by the arrows is the reconstructed area.

Multi-directional scanning: Due to the inherent limitations of the framework, we note that rapid movement may lead to the omission of some spectra (see the Applicability and limitation section). A solution lies in simultaneous multi-directional scanning using multiple scanning spectrometers. For instance, spectrometers scanning in opposite directions (left to right and right to left) ensure that horizontally moving objects consistently meet the counter-motion conditions, as discussed in the Applicability and limitation section. This ensures accurate reconstruction to reduce spectrum omission caused by too fast movement. In addition, an extra HSI sensor is available to capture spectral information from obscured areas that cannot be acquired by a single HSI sensor. Multi-directional scanning enhances the system's ability to capture and reconstruct motion with improved precision and reliability.

Point-scan hyperspectral imaging: Based on the reported *SpectraTrack* framework, point-scan hyperspectral cameras can also accomplish video imaging that is challenging to accomplish using conventional methods. Displacements or vibrations within the FOV can introduce artifacts or overlaps. Our proposed method addresses these issues by estimating target motion, thus enhancing imaging speed and quality. By mitigating issues related to scanning time and artifacts, point-scan hyperspectral cameras are capable of enabling more efficient and accurate video imaging.

More accurate motion tracking: The accuracy of motion flow estimation is crucial for the quality of reconstruction in our system. Integrating deep learning for dense motion tracking⁴³ and rendering could yield superior reconstruction results. Additionally, event cameras⁴⁴ offer high temporal resolution and low latency, making them suitable for capturing fast motion. Thus, replacing the

conventional RGB camera with event cameras for motion information acquisition holds the potential for further quality enhancement.

Variations in spectral responses of target materials and the illumination: This technique is built upon the assumption of no variations in spectral responses of target materials and the illumination. For scenes that do not adhere to this assumption, one approach is to establish a correspondence table between time and material spectral variations. In addition, we can estimate scene illumination changes through the RGB camera or other sensors. Finally, reconstructing the HSI images corresponding to each moment can be achieved through optimization or deep learning methods based on the correspondence table and the captured illumination changes.

Reducing redundancy: One major redundancy is that the system provides up to 1200 channels between 400 and 1000 nm with 2.5 nm spectral resolution, leading to redundancy in the spectral dimension. In general, fewer channels (e.g., 300 channels) can meet the requirements for common applications. To achieve higher spectral resolution, we can apply spectral super-resolution methods^{34,45} to decouple more precise spectral information. Another issue is the storage redundancy. The *SpectraTrack* technique involves reconstructing large volumes of data. These reconstructions contain substantial redundant data. For classification and detection, we can employ band selection methods to choose specific spectral bands for post-processing⁴⁶. Hyperspectral image compression⁴⁷ is another alternative to ease the storage burden. What's more, for drone-based experiments, the scene's motion can be described by a set of motion parameters instead of optical flows. Storing these parameters along with the line-scan measurement would then enable direct recovery through the *SpectraTrack* reconstruction algorithm, significantly reducing the storage burden.

Higher integration of multiple imaging sensors: Drawing inspiration from multi-camera cell phone systems, we envision developing a low-cost, highly integrated, high-resolution HSI system. This could be achieved by miniaturizing compact spectrometer^{48,49} and scanning devices, making the system more compact and portable. By incorporating a multi-camera setup, the system can capture and process data from multiple viewpoints simultaneously, enabling advanced imaging techniques such as plenoptic imaging⁵⁰. This higher level of integration expands the possibilities for applications that require detailed and comprehensive imaging capabilities.

Methods

Spatio-temporal-spectral multiplexing of scanning-based imaging spectrometers

Given the 4D spatio-temporal-spectral data $\mathbf{D}_{T,X,Y,C}$ of target scene, the 3D measurement \mathbf{H} of the line-scan spectrometer is denoted as

$$\mathbf{H}_{X,Y,C} = \mathcal{S}(\mathbf{D}_{T,X,Y,C}), \quad (1)$$

where T, X, Y, C are the resolutions of temporal (t), spatial (x, y), and spectral (λ) domain respectively. The operator \mathcal{S} represents the scanning process that records the hyperspectral data. Given the line-scan spectrometer scans along the x -dimension,

$$\mathbf{H}(:, y, \lambda) = \mathcal{S}(\mathbf{D}(:, :, y, \lambda)) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}_{T,X} \otimes \mathbf{D}(:, :, y, \lambda), \quad (2)$$

where ‘ \otimes ’ denotes the hadamard product and ‘ \cdot ’ stands for all the elements in this dimension. Here we consider a single scanning process, thereby $T = X = N$. As illustrated in Fig. 1A, the measurement H compresses both the spatial (x) and temporal (t) information of original 4D hyperspectral data into one dimension. Each 2D slice $\mathbf{H}(n, :, :)$ corresponds to $\mathbf{D}(n, n, :, :)$ which captured at n -th acquisition timing of once scanning. Under the assumption that there is no illumination and spectral reflectance change for the target during scanning, we can deduce the spatial movement of each point in H to backfill the uncaptured positions in $\mathbf{D}_{T,X,Y,C}$ by auxiliary optical flow introduced by a synchronous array camera (Fig. 2).

Motion-guided reconstruction

When the target object is moving during scanning, the spectral response $\mathbf{H}(n, y, :)$ at time n and position (n, y) will appear at other times and at other positions such as time $n + dt$ and position $(n + dx, y + dy)$. Therefore, we can backfill the uncaptured spectral response $\mathbf{D}(n + dt, n + dx, y + dy, :) = \mathbf{H}(n, y, :)$. The movement of each point between two times can be obtained by the dense optical flow extracted by the RGB video $\mathbf{G}_{T,X,Y}$:

$$\begin{aligned} \Delta_{n \rightarrow n+dt}^h, \Delta_{n \rightarrow n+dt}^v &= \text{OF}(\mathbf{G}(n, :, :), \mathbf{G}(n + dt, :, :)), \\ dx, dy &= \Delta_{n \rightarrow n+dt}^h(x, y), \Delta_{n \rightarrow n+dt}^v(x, y), \end{aligned} \quad (3)$$

where OF stands for the dense optical flow operator, $\Delta_{n \rightarrow n+dt}^h$ and $\Delta_{n \rightarrow n+dt}^v$ are the horizontal and vertical dense optical flows, respectively. One can repeat the above steps to backfill spectra in $\mathbf{D}_{T,X,Y,C}$ of all the times in one scanning. In the experiments, we applied the DeepFlow⁵¹ method for optical flow extraction.

Due to the decreased accuracy of optical flow techniques under high motion, we developed an incremental transformation method. This approach calculates the coordinate transformation between frames that are distant from each other by utilizing optical flow from

multiple adjacent frames. Suppose that there are multiple intermediate frames, $\mathbf{G}_1, \mathbf{G}_2, \dots, \mathbf{G}_{n-1}$, between two frames, \mathbf{G}_0 and \mathbf{G}_n . For simple uniformly rigid motion scenes, the optical flow between \mathbf{G}_0 and \mathbf{G}_n can be represented as the summation of optical flows between all frames from \mathbf{G}_0 to \mathbf{G}_n .

$$\text{OF}(\mathbf{G}_0, \mathbf{G}_n) = \text{OF}(\mathbf{G}_0, \mathbf{G}_1) + \text{OF}(\mathbf{G}_1, \mathbf{G}_2) + \dots + \text{OF}(\mathbf{G}_{n-1}, \mathbf{G}_n) \quad (4)$$

Furthermore, if the target motion does not conform to the aforementioned assumption, we calculate the incremental coordinate transformations using the optical flows between adjacent frames and obtain the high motion coordinate transformation as follows.

$$\Delta_{0 \rightarrow n}^h, \Delta_{0 \rightarrow n}^v = \mathcal{T}(\mathbf{G}_0, \mathbf{G}_n) = \mathcal{T}(\dots \mathcal{T}(\mathcal{T}(\mathbf{G}_0, \mathbf{G}_1), \mathbf{G}_2), \dots, \mathbf{G}_n), \quad (5)$$

where \mathcal{T} denotes the coordinate transformation operator. Specifically, it refers to the `grid_sample` function in Pytorch which warps images with the dense optical flow. More details of the reconstruction algorithm can be found in Supplementary Note 2–4.

The reconstructed data $\mathbf{D}_{T,X,Y,C}$ has $T = X$ scanning times, Y pixels per scanned line, and C spectral channels. The value of T or X is determined by the scanning speed and range while Y is determined by the spatial resolution of the imaging spectrometer. The *SpectraTrack* method reintegrates the spectra as a whole to reconstruct HSI videos. Therefore the spectral resolution C is the same as that of the imaging spectrometer used.

The authors affirm that human research participants provided informed consent for publication of the images in Fig. 5.

Data availability

The data that support the findings of this study is available under restricted access following the funded project requirements. Access can be obtained upon request to the corresponding author.

Code availability

The demo code of the reported technique is available under restricted access following the funded project requirements. Access can be obtained upon request to the corresponding author.

References

- Susarla, S. et al. Hyperspectral imaging of exciton confinement within a moiré unit cell with a subnanometer electron probe. *Science* **378**, 1235–1239 (2022).
- Bhargava, R. & Falahkheirkhah, K. Enhancing hyperspectral imaging. *Nat. Mach. Intell.* **3**, 279–280 (2021).
- Brady, D. J. *Optical imaging and spectroscopy* (John Wiley & Sons, 2009).
- Valm, A. M. et al. Applying systems-level spectral imaging and analysis to reveal the organelle interactome. *Nature* **546**, 162–167 (2017).
- Ozeki, Y. et al. High-speed molecular spectral imaging of tissue with stimulated raman scattering. *Nat. Photon.* **6**, 845–851 (2012).
- Groner, W. et al. Orthogonal polarization spectral imaging: a new method for study of the microcirculation. *Nat. Med.* **5**, 1209–1212 (1999).
- Cutrale, F. et al. Hyperspectral phasor analysis enables multiplexed 5d in vivo imaging. *Nat. Methods* **14**, 149–152 (2017).
- Backman, V. et al. Detection of preinvasive cancer cells. *Nature* **406**, 35–36 (2000).
- Li, Q. et al. Review of spectral imaging technology in biomedical engineering: achievements and challenges. *J. Biomed. Opt.* **18**, 100901–100901 (2013).
- Mason, R. E. et al. Evidence, causes, and consequences of declining nitrogen availability in terrestrial ecosystems. *Science* **376**, eabh3767 (2022).

11. Lu, B., Dao, P. D., Liu, J., He, Y. & Shang, J. Recent advances of hyperspectral imaging technology and applications in agriculture. *Remote Sens.* **12**, 2659 (2020).
12. Wang, C. et al. A review of deep learning used in the hyperspectral image analysis for agriculture. *Artif. Intell. Rev.* **54**, 5205–5253 (2021).
13. Bao, F. et al. Heat-assisted detection and ranging. *Nature* **619**, 743–748 (2023).
14. Sumriddetchkajorn, S. & Intaravanne, Y. Hyperspectral imaging-based credit card verifier structure with adaptive learning. *Appl. Opt.* **47**, 6594–6600 (2008).
15. Rickard, L. J., Basedow, R. W., Zalewski, E. F., Silvergate, P. R. & Landers, M. HYDICE: An airborne system for hyperspectral imaging. *Imag. Spectrom. Terrestrial Environ.* **1937**, 173–179 (1993).
16. Basedow, R. W., Carmer, D. C. & Anderson, M. E. HYDICE system: Implementation and performance. *Imag. Spectrom.* **2480**, 258–267 (1995).
17. Gupta, N., Dahmani, R. & Choy, S. Acousto-optic tunable filter based visible-to near-infrared spectropolarimetric imager. *Opt. Eng.* **41**, 1033–1038 (2002).
18. Gupta, N. Hyperspectral imager development at army research laboratory. *Infrared Technol. Appl.* XXXIV, **6940**, 573–582 (2008).
19. Fang, J. et al. Wide-field mid-infrared hyperspectral imaging beyond video rate. *Nat. Commun.* **15**, 1811 (2024).
20. Cao, X., Du, H., Tong, X., Dai, Q. & Lin, S. A prism-mask system for multispectral video acquisition. *IEEE Trans. Pattern Anal. Mach. Intell.* **33**, 2423–2435 (2011).
21. Mu, T., Han, F., Bao, D., Zhang, C. & Liang, R. Compact snapshot optically replicating and remapping imaging spectrometer (orris) using a focal plane continuous variable filter. *Opt. Lett.* **44**, 1281–1284 (2019).
22. Park, J., Feng, X., Liang, R. & Gao, L. Snapshot multidimensional photography through active optical mapping. *Nat. Comm.* **11**, 5602 (2020).
23. McClung, A., Samudrala, S., Torfeh, M., Mansouree, M. & Arbabi, A. Snapshot spectral imaging with parallel metasystems. *Sci. Adv.* **6**, eabc7646 (2020).
24. Yako, M. et al. Video-rate hyperspectral camera based on a cmos-compatible random array of fabry–pérot filters. *Nat. Photon.* **17**, 218–223 (2023).
25. Descour, M. & Dereniak, E. Computed-tomography imaging spectrometer: experimental calibration and reconstruction results. *Appl. Opt.* **34**, 4817–4826 (1995).
26. Yuan, L., Song, Q., Liu, H., Heggarty, K. & Cai, W. Super-resolution computed tomography imaging spectrometry. *Photonics Res.* **11**, 212–224 (2023).
27. Gehm, M. E., John, R., Brady, D. J., Willett, R. M. & Schulz, T. J. Single-shot compressive spectral imaging with a dual-disperser architecture. *Opt. Express* **15**, 14013–14027 (2007).
28. Wagadarikar, A., John, R., Willett, R. & Brady, D. Single disperser design for coded aperture snapshot spectral imaging. *Appl. Opt.* **47**, B44–B51 (2008).
29. Arguello, H. & Arce, G. R. Colored coded aperture design by concentration of measure in compressive spectral imaging. *IEEE Trans. Image Process.* **23**, 1896–1908 (2014).
30. Lin, X., Liu, Y., Wu, J. & Dai, Q. Spatial-spectral encoded compressive hyperspectral imaging. *ACM Trans. Graph.* **33**, 1–11 (2014).
31. Xu, Y., Lu, L., Saragadam, V. & Kelly, K. F. A compressive hyperspectral video imaging system using a single-pixel detector. *Nat. Comm.* **15**, 1456 (2024).
32. Jeon, D. S. et al. Compact snapshot hyperspectral imaging with diffracted rotation. *ACM Trans. Graph.* **38**, 1–13 (2019).
33. Monakhova, K., Yanny, K., Aggarwal, N. & Waller, L. Spectral diffusercam: lensless snapshot hyperspectral imaging with a spectral filter array. *Optica* **7**, 1298–1307 (2020).
34. Xiong, Z. et al. HSCNN: Cnn-based hyperspectral image recovery from spectrally undersampled projections. In *Proceedings IEEE International Conference on Computer Vision Workshops (ICCVW)* 518–525 (2017).
35. Nie, S. et al. Deeply learned filter response functions for hyperspectral reconstruction. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 4767–4776 (2018).
36. Wang, L., Zhang, T., Fu, Y. & Huang, H. HyperReconNet: Joint coded aperture optimization and image reconstruction for compressive hyperspectral imaging. *IEEE Trans. Image Process.* **28**, 2257–2270 (2018).
37. Miao, X., Yuan, X., Pu, Y. & Athitsos, V. λ -net: Reconstruct hyperspectral images from a snapshot measurement. In *IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, 4059–4069 (2019).
38. Huang, L., Luo, R., Liu, X. & Hao, X. Spectral imaging with deep learning. *Light Sci. Appl.* **11**, 61 (2022).
39. Gat, N. Imaging spectroscopy using tunable filters: a review. *Wavel. Appl. VII* **4056**, 50–64 (2000).
40. Tong, Q., Xue, Y. & Zhang, L. Progress in hyperspectral remote sensing science and technology in china over the past three decades. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **7**, 70–91 (2013).
41. Liu, S., Yang, B., Yuen, P. C. & Zhao, G. A 3d mask face anti-spoofing database with real-world variations. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 100–106 (IEEE, 2016).
42. Rao, S., Huang, Y., Cui, K. & Li, Y. Anti-spoofing face recognition using a metasurface-based snapshot hyperspectral image sensor. *Optica* **9**, 1253–1259 (2022).
43. Wang, Q. et al. Tracking everything everywhere all at once. In *Proceedings of IEEE International Conference on Computer Vision (IEEE, 2023)*.
44. Rebecq, H., Ranftl, R., Koltun, V. & Scaramuzza, D. High speed and high dynamic range video with an event camera. *IEEE Trans. Pattern Anal. Mach. Intell.* **43**, 1964–1980 (2019).
45. Yi, C., Zhao, Y.-Q. & Chan, J. C.-W. Spectral super-resolution for multispectral image based on spectral improvement strategy and spatial preservation strategy. *IEEE Trans. Geosci. Remote Sens.* **57**, 9010–9024 (2019).
46. Sun, W. & Du., Q. Hyperspectral band selection: A review. *IEEE Geosci. Remote Sens. Mag.* **7**, 118–139 (2019).
47. Dua, Y., Kumar, V. & Singh, R. S. Comprehensive review of hyperspectral image compression algorithms. *Opt. Eng.* **59**, 090902 (2020).
48. Yang, Z. et al. Single-nanowire spectrometers. *Science* **365**, 1017–1020 (2019).
49. Yang, Z., Albrow-Owen, T., Cai, W. & Hasan, T. Miniaturization of optical spectrometers. *Science* **371**, eabe0722 (2021).
50. Lu, Z. et al. Virtual-scanning light-field microscopy for robust snapshot high-resolution volumetric imaging. *Nat. Methods* **20**, 735–746 (2023).
51. Weinzaepfel, P., Revaud, J., Harchaoui, Z. & Schmid, C. DeepFlow: Large displacement optical flow with deep matching. In *Proceedings of IEEE International Conference on Computer Vision*, 1385–1392 (IEEE, 2013).

Acknowledgements

This work was supported by the National Natural Science Foundation of China under Grants 62322502 (L.B.), 61827901 (L.B.), and 62088101 (L.B.), and the Guangdong Province Key Laboratory of Intelligent Detection in Complex Environment of Aerospace, Land and Sea under Grant 2022KSYS016 (L.B.). We thank FigSpec Co., Ltd. in China for providing help in setting up the system optics and the HSI measurement under high winds (Fig. 4A).

Author contributions

D.L. and L.B. conceived the idea. D.L. and J.Z. built the prototype system. D.L. and J.W. developed the codes for the reconstruction algorithm. J.Z. developed the codes for hardware control. D.L., J.W., J.Z., and H.X. conducted the experiments. L.B. supervised the project. All the authors were involved in discussions and contributed to writing and revising the manuscript during the project.

Competing interests

L.B. and D.L. hold patents on technologies related to the systems developed in this work (China patent numbers ZL202210764141.5, ZL202310982648.2, and ZL202311052634.7) and submitted related patent applications. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-53747-8>.

Correspondence and requests for materials should be addressed to Liheng Bian.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024