remote sensing

MDPI

Article

Efficient Hyperspectral Video Reconstruction via Dual-Channel DMD Encoding

Mingming Ma [1], Yi Niu [1,2,*], Dahua Gao [1], Fu Li [1] and Guangming Shi [1,2]

[1] School of Artificial Intelligence, Xidian University, Xi'an 710071, China; mamm@stu.xidian.edu.cn (M.M.); dhgao@xidian.edu.cn (D.G.); fuli@mail.xidian.edu.cn (F.L.); gmshi@xidian.edu.cn (G.S.)
[2] Peng Cheng Laboratory, Shenzhen 518055, China
* Correspondence: niuyi@mail.xidian.edu.cn

Abstract: Hyperspectral video acquisition requires a precise balance between spectral and temporal resolution, often achieved through compressive sampling using two-dimensional detectors and spectral reconstruction algorithms. However, the reliance on spatial light modulators for coding reduces optical efficiency, while complex recovery algorithms hinder real-time reconstruction. To address these challenges, we propose a digital-micromirror-device-based complementary dual-channel hyperspectral (DMD-CDH) video imaging system. This system employs a DMD for simultaneous light splitting and spatial encoding, enabling one channel to perform non-aliasing spectral sampling at lower frame rates while the other provides complementary high-rate sampling for panchromatic video. Featuring high optical throughput and efficient complementary sampling, the system ensures reliable hyperspectral video reconstruction and serves as a robust ground-based validation platform for remote sensing applications. Additionally, we introduce tailored optical error calibration and fixation techniques alongside a lightweight hyperspectral fusion network for reconstruction, achieving hyperspectral frame rates exceeding 30 fps. Compared to the existing models, this system simplifies the calibration process and provides a practical high-performance solution for real-time hyperspectral video imaging.

Keywords: hyperspectral image; spectral video; digital micromirror device (DMD); fusion network

check for updates

Academic Editor: Salah Bourennane

Received: 23 November 2024
Revised: 25 December 2024
Accepted: 1 January 2025
Published: 8 January 2025

Citation: Ma, M.; Niu, Y.; Gao, D.; Li, F.; Shi, G. Efficient Hyperspectral Video Reconstruction via Dual-Channel DMD Encoding. *Remote Sens.* **2025**, *17*, 190. https://doi.org/10.3390/rs17020190

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Hyperspectral imaging (HSI) technology is a photoelectric detection and recognition technology that comprises the combination of spectroscopy and optical imaging. This cutting-edge technology simultaneously captures both spatial images and spectral information of objects, effectively integrating imagery with spectrum analysis. The spectral images produced offer intricate details of scenes, making HSI a valuable tool in diverse applications like remote sensing, computer vision [1,2], detection [3,4], and medical diagnostics. However, the traditional hyperspectral imaging systems encounter challenges, such as limited light throughput and lengthy scanning processes to acquire a three-dimensional data cube, rendering the capture of dynamic scenes challenging. Established techniques like whiskbroom [5] or pushbroom [6] spatial scanning, as well as spectral-scan-based tunable filters [7,8] or filter wheels, necessitate multiple exposures across spatial or spectral dimensions, compromising temporal sampling rates. Consequently, these methods are predominantly suitable for static or moderately dynamic scenes. Recent years have witnessed significant strides in snapshot HSI technology, revolutionizing the field. This innovative approach captures all spectral and spatial information in a single exposure, eliminating the

need for scanning and enabling the real-time capture of dynamic scenes. This breakthrough has expanded the horizons of HSI applications [9–11]. In the field of computer vision, it has facilitated region analysis through invariant color representations [12], and, in remote sensing, it has enhanced network capabilities for similar analyses [13,14]. Additionally, detection efficiencies have undergone substantial improvements with the integration of HSI [15,16], paving the way for new frontiers in hyperspectral imaging applications.

Hyperspectral imaging of dynamic scenes, which captures extensive data across the spatial, spectral, and temporal domains, has evolved significantly over recent decades. The characterization of such imaging systems involves four key parameters: $\mathbf{N_x}$, $\mathbf{N_y}$, $\mathbf{N_c}$ and $\mathbf{N_t}$, collectively forming the "hypercube" framework [17]. This shift towards hyperspectral video imaging, propelled by advancements in computational imaging, marks a departure from traditional still image capture. Computational spectral imaging eliminates the need for scanning operations by using a secondary imaging system. It effectively projects the spatially and spectrally encoded 3D data cube onto a plane, overlaying the aliased images of detected spatial and spectral information. The original data are then computationally restored, simultaneously retrieving information in both the spatial and spectral dimensions. Techniques like Computed Tomographic Imaging Spectrometry (CTIS) [18], Coded Aperture Snapshot Spectral Imager (CASSI) [19,20], and Spatial–Spectral Compressive Spectral Imaging (SSCSI) [20–22] utilize novel optical designs for compressive sampling with a 2D detector. However, these methods often sacrifice spatial and spectral resolutions and are not ideal for real-time applications due to complex processing needs. Recent approaches, like compact spectral imaging using prisms [23,24] or diffractive snapshot hyperspectral imaging with diffractive optical elements (DOEs) [25–27], aim to simplify spectral imaging systems. These methods focus on using advanced primary imaging devices and minimizing the need for secondary imaging, but they still face challenges in coding flexibility and reconstruction of small objects. Therefore, our work concentrates on researching and optimizing secondary imaging systems to minimize assembly complexity, solve calibration issues, and enable efficient coding of small moving targets, paving the way for practical applications of this technology.

To date, our research has primarily concentrated on hyperspectral video imagers capable of single-snapshot measurements. In this context, we have identified and systematically summarized the four principal limitations inherent to the classical CASSI technology. Additionally, we have meticulously evaluated various methodological advancements aimed at overcoming these shortcomings, thereby contributing to the refinement and enhancement of hyperspectral imaging technology.

(1) High computational complexity: This is caused by the complicated optimization for compressive sensing restoration.
(2) Lack of online prior: The CASSI restoration is an ill-posed inverse problem and relies on prior models, like TV (Total Variation) [28] and DCT (Discrete Cosine Transform) [29], to regularize the optimization. However, the handcrafted prior models can only characterize the offline general features of the multispectral images but fail to characterize the online features of the scene.
(3) Low optical efficiency: Because of the spatial coding, the incident light has to be blocked by the template, which significantly decreases the optical efficiency of the system.
(4) Mechanical coding template: The original coding template for the CASSI system is created via a photolithography mask, which cannot be changed according to the scene.

In an effort to reduce the complexity of spectral reconstruction algorithms, deep-learning-based methodologies [30–33] have been increasingly adopted. These approaches serve as sophisticated alternatives to traditional reconstruction processes and have markedly advanced hyperspectral fusion network capabilities. The integration of RGB cameras in the complementary dual-channel CASSI system also provides innovative strate-

gies for augmenting prior information. Cao et al. [34,35] innovatively proposed a Hybrid Spectral Video Imaging System (HVIS), employing a beam splitter to bifurcate the incident light into dual channels. This setup enables one channel to capture the scene's spectral information via a sparse non-overlapping coding template while the other channel utilizes a standard RGB camera for online capture preceding the reconstruction phase [36]. This technique strategically bypasses the need for complex recursive optimization, favoring more streamlined bilateral/trilateral filtering methods. Nonetheless, the direct imaging application of beam splitters and sparse templates often incurs significant optical efficiency losses, thus posing challenges to the HVIS technique's adaptability in terms of template modifications and further exacerbating light efficiency issues. As a viable alternative, traditional mechanical coding templates [37,38] may be effectively substituted with digital micromirror devices [39–41] or liquid crystal spatial light modulators (LC-SLMs) [42–44], thereby aiding in encoding pattern adjustments. Despite the potential of adaptive coding patterns to elevate imaging quality, the issue of light efficiency continues to be a pertinent concern. Broadly, the design of hardware spectral imaging systems must prioritize high light efficiency to ensure elevated frame imaging rates while simultaneously striving to reduce the computational complexity of reconstruction algorithms, a critical factor for enabling real-time spectral image acquisition.

In this paper, we propose a novel methodology designed to address the quartet of challenges presently impeding the progress of hyperspectral video imaging. Our approach is anchored in a novel digital micromirror device (DMD)-based complementary dual-channel hyperspectral video imaging system. This system stands out for its ability to achieve simultaneous light splitting and spatial encoding, leveraging the high light efficiency characteristic of the DMD device. A pivotal feature of our DMD-CDH system is its capacity to generate side-band information through complementary reflections of the DMD device, which serves as an invaluable online prior during spatial encoding. This feature is especially significant for our system's simplistic yet robust geometric calibration process, which is essential for pixel-level spectral recovery between the spatially encoded sensor and the online prior channel sensor. Our investigation also delves into crucial practical considerations, such as addressing diamond–square matching and filling gaps in the online prior channel. To complete the spectral reconstruction, we employ a lightweight hyperspectral fusion network. The capabilities of our system have the potential to be extended to validate satellite remote sensing techniques and facilitate real-time acquisition of spectral information from target scenes.

The proposed system is inspired by the limitations of the current hyperspectral imaging techniques, including low optical efficiency, complex calibration requirements, and the trade-off between spectral and temporal resolution. By leveraging the high reflectance and spatial encoding capabilities of the DMD, we aim to create a system that efficiently addresses these challenges while enabling real-time spectral reconstruction.

In summary, the main contributions are summarized as follows.

(1) Innovative dual-channel hyperspectral video imaging system: Overcoming high computational complexity and other traditional challenges, this novel system with a DMD device enables real-time spectral data acquisition, offering more efficient hyperspectral video imaging.
(2) Advanced imaging techniques and flexibility: By supporting scene-encoded snapshot imaging at video rates and pushbroom mode, along with quickly switchable coding templates, the system significantly enhances its versatility and adaptability.
(3) Addressing optical and engineering challenges: Through comprehensive analysis and development of a new calibration method, this research tackles the diamond–

square matching issue and enhances system efficiency, emphasizing the importance of practical engineering considerations in hyperspectral imaging.

(4) Integration of lightweight convolutional neural network for real-time processing: Employing a neural network for the fusion of hyperspectral and color images, this approach greatly improves the system's accuracy and efficiency, streamlining spectral data capturing.

The remaining sections of this paper are organized as follows. Section 2 provides a detailed description of our proposed framework, outlining its key features and highlighting its advantages over the traditional hyperspectral imaging systems. In Section 3, we present comparative experiments that demonstrate the effectiveness and efficiency of our proposed approach and discuss the experimental results to further validate our system's performance. Finally, in Section 4, we draw conclusions regarding the potential impact of our approach in the field.

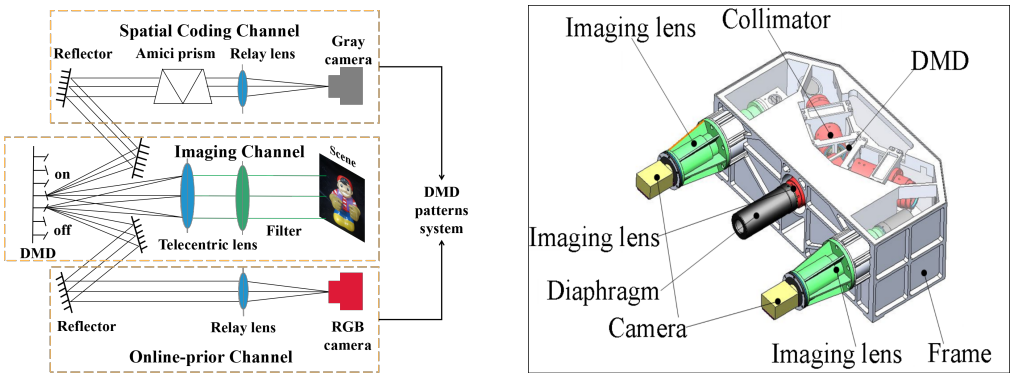
2. Proposed Method

In this section, we will provide a detailed introduction to the principle and design of our proposed computational spectral imaging system. We will then analyze the pixel-level registration problem associated with the system and discuss potential solutions to this challenge. Additionally, we will detail the methods we used to complete the multispectral image reconstruction task, including the process of filling online channel pixels and implementation of a convolutional neural network.

2.1. Principle and Design

The DMD-CDH system introduced in this paper is a symmetrical three-channel system, comprising an imaging channel, a spatial coding channel, and an online prior channel, as illustrated in Figure 1. At the heart of this system lies the imaging channel, which captures light from the scene. This light traverses an optical assembly consisting of a telecentric lens and a visible light filter, operational within the 380–900 nm transmission window. The light subsequently forms optical images on a Texas Instruments 0.45 WXGA DMD chipset (manufactured by Texas Instruments, Dallas, TX, USA), which plays a pivotal role in spatially coding the light by manipulating each micromirror with a bespoke encoding template. The light, thus modulated, is then strategically redirected into either the spatial encoding channel or the online prior channel via a dual-mirror setup designed to alter the optical path and circumvent hardware collisions inherent in direct imaging. In the spatial coding channel, an off-the-shelf Amici prism disperses the reflected light, creating spectrally dispersed images of the DMD-modulated image intensity with specific templates on a grayscale video camera. The spectral resolution here hinges on the Amici prism's design, while the dispersion orientation is defined by its placement. Conversely, the online prior channel employs a relay lens to modify the focus, thereby generating spatial color images of the DMD-modulated intensity with complementary templates on an RGB video camera—a process mirroring that of the spatial encoding channel.

This innovative system, adept at self-calibration, concurrently acquires low-resolution hyperspectral video frames and high-resolution RGB video frames. This dual acquisition mechanism greatly enhances the efficiency and accuracy of spectral and spatial information gathering. By leveraging this advanced three-channel system, we achieve high light efficiency and adaptable template use, enabling real-time capture of spectral data from target scenes with significantly reduced computational complexity.



(**a**) (**b**)

Figure 1. (**a**) Schematic of the DMD-CDM system. (**b**) Imaging system design diagram with stable structure.

The DMD employed in our multispectral imaging system serves as a highly efficient spatial light modulator, offering precise and rapid control of light reflection. As depicted in Figure 2, the DMD utilized in our setup features a diamond-shaped arrangement, consisting of 912 × 1140 aluminum micromirrors, each with a pixel pitch of 7.6 µm. These micromirrors are individually addressable and capable of tilting about their diagonal axes, enabling dynamic light modulation with exceptional precision.

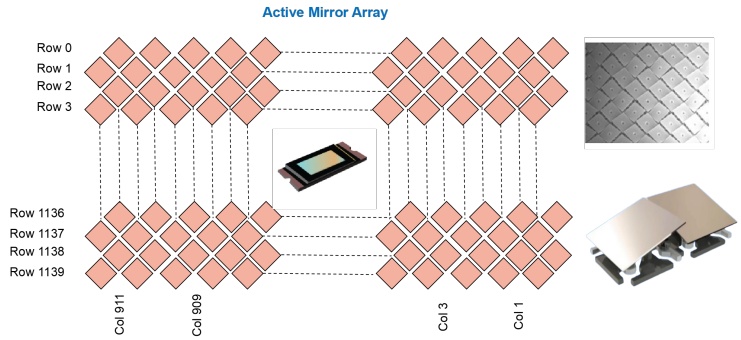


Figure 2. Schematic diagram of DMD structure, with diamond-shaped arrangement reversed diagonally to both sides at a specified angle.

Structurally, the DMD surface is composed of thousands of reflective micromirrors, each representing a single pixel. The micromirrors are mounted on torsion hinges composed of aluminum, which allow them to tilt at three discrete states: $+12°$, $0°$, and $-12°$. These tilt angles are achieved through electrostatic actuation, where a voltage difference is applied to control electrodes located beneath the micromirror. The applied voltage creates electrostatic forces that cause the mirror to pivot about its hinge. To ensure stability, the system utilizes damping springs as contact points, which act as mechanical stops to prevent overshooting and enable precise positioning of the micromirrors. Each micromirror is separated by a gap of approximately 1 µm, minimizing crosstalk between adjacent mirrors and ensuring high optical efficiency.

In our system, micromirrors in the $+12°$ ('on') state reflect light toward the online prior channel, while those in the $-12°$ ('off') state deflect light toward the spatial coding channel. The $0°$ state is a transitional position that occurs during switching. The rapid switching capability of the micromirrors, with a refresh rate exceeding 4000 Hz, allows the DMD to dynamically generate complex spatial light patterns in real time, which is critical for high-throughput hyperspectral imaging.

Additionally, the micromirrors are designed for durability, capable of withstanding over one million actuation cycles without degradation. This makes the DMD a reliable and robust solution for applications requiring continuous operation. By modulating the tilt angles of individual micromirrors, the DMD controls the direction of reflected light, enabling precise-intensity modulation at the pixel level. This functionality translates directly to the brightness or darkness of corresponding pixels in the generated digital image.

To validate the feasibility of our innovative imaging system, we engineered a theoretical prototype in a laboratory setting and further integrated it with a novel three-arm spectral imaging system configuration, as illustrated in Figure 1b. The calibration process was meticulously carried out using a laser to determine the optimal deflection angles of both mirror sets. This precision was complemented by subsequent fine-tuning and secure placement of the dispersive prisms and detectors. A critical design aspect involved ensuring that the micromirrors on the diamond-shaped DMD were accurately deflected either to the left or right along their diagonal axis, thus preventing any disparity in height across the two reflective channels. Emphasizing accuracy and reliability, we anchored the entire imaging system frame to be fixed and structurally robust. The integration of the designed spectral imaging system with the three-arm structure facilitates real-time imaging of scenes, demonstrating the potential for practical applications of our proposed system.

To provide a clearer overview of the technical specifications and performance of the proposed DMD-CDH system, Table 1 summarizes the key parameters. The system covers a spectral range from 400 to 900 nm with 30 spectral channels and achieves a spatial resolution of 1140 × 912 pixels and a spectral resolution of 12.5 nm. With a compact design, the system weighs approximately 15 kg and has dimensions of 50.0 × 50.0 × 18.0 cm. The connection interface is USB 3.0, and the pixel size of the detector is 7.6 µm. Additionally, the DMD operates with a tilt angle of ±12°. These specifications highlight the system's compactness and versatility for real-time hyperspectral imaging tasks, providing a robust foundation for further evaluation and application in various scenarios.

Table 1. DMD-CDH system design specifications and performance.

Parameter	Value
Spectral range (nm)	400–900
Number of spectral channels	30
Spatial resolution	1140 × 912
Spectral resolution	12.5
Weight (kg)	15
Size (cm × cm × cm)	50.0 × 50.0 × 18.0
Connection method	USB 3.0
Pixel size (µm)	7.6
Tilt angle (○)	±12

Spectral imaging systems based on beam splitters will lead to a reduction in optical efficiency [39]; in stark contrast, our novel DMD-CDH video imaging system circumvents these drawbacks. It obviates the need for a beam splitter and is adept at harnessing all incident light energy for both the spatial coding and the online prior channels. A cornerstone of our system is the high-accuracy pixel-level geometric calibration process, essential for spectral restoration between the spatial coding channel sensor and the online prior channel sensor. Unlike beam-splitter-based systems, where calibration can be arduous due to noise and optical system aberrations, our DMD-CDH system simplifies this process. Both channels in DMD-CDH are secondary imageries of the DMD plane, with each micromirror acting as a distinct feature point across both sensors. Calibration becomes straightforward, achieved by simply flipping the micromirrors twice to reflect uniform incident light into both channels. With

these advancements, the DMD-CDH system offers a more flexible and efficient hyperspectral imaging solution that could have a wide range of practical applications.

2.2. Diamond–Square Matching Problem

Current DMD systems feature rectangular target surfaces composed of diamond-shaped cells, with each micromirror capable of flipping only along its diagonal. To align these 'square-to-square-mirror-pixels' with the sensor pixels, standard practices necessitate rotating the DMD or tilting the camera by 45° [38,39]. This approach, however, can result in pixel waste and complicates the system's assembly and commissioning processes. As depicted in Figure 3, a significant proportion of the micromirrors cannot be effectively captured by the sensor without employing a larger sensor. While the rotation of the DMD or the camera tilt might aid algorithmic simulation, these adaptations are suboptimal for practical deployment. Addressing these challenges is crucial for the advancement of DMD-based hyperspectral imaging systems, necessitating further development to optimize their design and functionality.

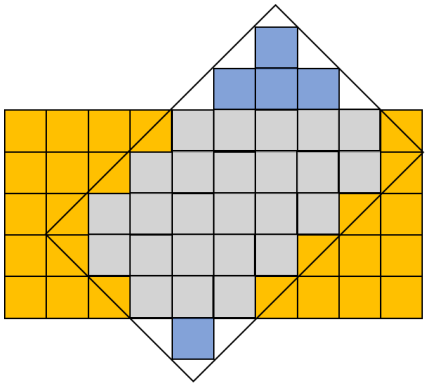


Figure 3. Loss of field of view due to rotating DMD. The yellow squares represent pixels on the detector, the gray squares represent DMD pixels aligned with the detector after tilting 45°, and the blue squares represent pixels that cannot be captured by the detector.

To improve the stability of the system without requiring camera rotation, we use a diamond-shaped DMD as a spatial light modulator, ensuring that the reflected light path and incident light path lie in the same horizontal plane. In the DMD-CDH system, the DMD serves as the reference plane for both cameras, enabling dual-camera registration by calibrating the corresponding relationship between the cameras and DMD micromirrors separately. The diamond-shaped staggered DMD is shown in Figure 4a, while the square-shaped arranged CMOS is depicted in Figure 4b. In our design, each pixel in the DMD pattern is represented by a 1-pixel × 1-pixel subset, which is imaged onto a 2-pixel × 2-pixel CMOS subset using a relay lens with an appropriate focal length to achieve a one-to-two pixel ratio at the pixel level. However, limitations in actual assembly and the use of a screw-fixed structure can cause the pixel level ratio of CMOS and DMD to deviate from the standard 2:1, leading to the diamond–square matching problem. In practice, the possible correspondence between CMOS and DMD pixels is illustrated in Figure 4c. This means that sub-pixel registration cannot be achieved at the engineering level, and offsets in any direction during pixel projection on the DMD to the detector can make it difficult to find the correspondence between the two coordinate systems. To address this problem, previous studies have proposed hexagonal-tiled elements in the coding aperture design [45,46] or super-resolution algorithms [47] to solve the coding element and detector matching problem. However, for DMD-based systems, diffraction effects [41] of the device itself and sub-pixel-level alignment pose additional engineering challenges that can affect image quality. Calibration is crucial for the functionality of a computational imaging system.

Therefore, we developed an automatic self-calibration method that traverses the DMD to generate the coordinate positions of each micromirror and the two cameras, which needs to be performed only once to establish subsequent imaging conditions.

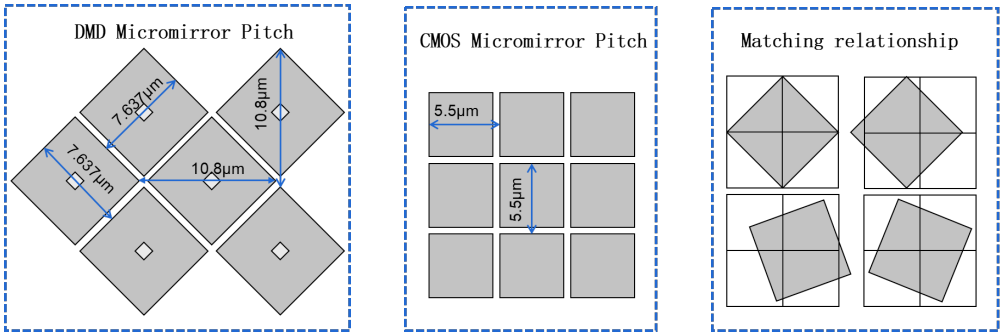


Figure 4. The arrangement and matching of DMD and detector. (**a**) Diamond-shaped staggered DMD. (**b**) Square-shaped arranged CMOS. (**c**) Schematic diagram of diamond–square matching problem.

Our calibration methodology for the DMD-CDH system is meticulously structured into a three-step process. Initially, we individually actuate each micromirror of the DMD to reflect through the online prior channel. During this stage, while subjected to white light illumination, we capture and record the specific coordinates of each point on the CMOS sensor. Subsequently, the process is repeated for the spatial coding channel, this time employing monochromatic light across a spectrum ranging from 450 nm to 900 nm, again recording the corresponding CMOS coordinates for each micromirror. The final step involves identifying the largest pixel correspondence for each point, culminating in the creation of two comprehensive comparison tables. These tables are instrumental in accurately mapping detector pixels back to their original source locations on the DMD. This calibration approach not only ensures precise alignment but also significantly enhances the overall quality and efficiency of the imaging results, optimizing the performance of our system.

Upon establishing the comparison table delineating the coordinate relationships between the two cameras and the DMD, our approach progresses to pushscan reconstruction utilizing lookup table matching. Opting for the largest pixel point as the micromirror response, however, tends to result in substantial CMOS point wastage and variable energy loss. This approach also introduces periodic blocking artifacts, as demonstrated in Figure 5a, due to the energy fluctuation with spatial positioning. To mitigate these issues, we implement second-order polynomial fitting as part of our calibration process. The least-squares method is employed to ascertain the polynomial coefficients, enabling us to accurately calculate both the coordinates and corresponding coefficients of the CMOS points associated with each DMD pixel. Subsequently, this information is used to update our lookup table, which now includes the two-pixel values and their respective coefficients for each DMD pixel across every spectral band. This updated lookup table is pivotal when encoding natural scenes with different templates as it enables precise indexing of the corresponding pixels for known points. This self-calibration method significantly bolsters the robustness of the DMD-CDH system for practical applications. Figure 5e,f showcases the improved results of the spectral image at 750 nm reconstructed using the newly derived lookup table, where the blocking artifact is notably minimized compared to the approach illustrated in Figure 5b,c that solely selects the energy maximum as the corresponding point.

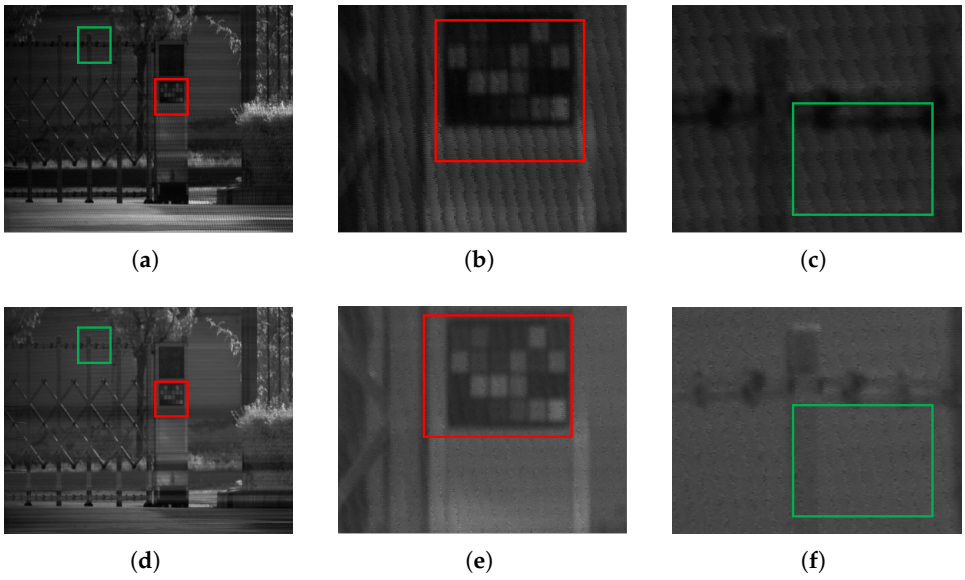


Figure 5. Reconstruction results in pushscan mode, the red and green boxes represent the selected partial regions. (**a**) Selecting the largest energy value to reconstruct the result. (**b**,**c**) Reconstruction detail display. (**d**) Polynomial fit calibration reconstruction result. (**e**,**f**) Reconstruction detail display.

2.3. Filling the Holes of the Online Prior Channel

The DMD-CDH system's capability extends beyond pushscan spectral reconstruction for static scenes. It can also perform real-time reconstruction of hyperspectral data for dynamic scenes using sparse sampling coding and recovery algorithms. Various hyperspectral imaging systems have different optical configurations and statistical properties of the spectral data. Different sampling strategies can lead to different sensing performances. Therefore, the sampling scheme of a hyperspectral acquisition system has a significant impact on the quality and speed of spectral reconstruction. However, the sampling is determined based on the optical and practical aspects of the spectrometer design. In this paper, we used a non-mixing undersampling coding template similar to the HVIS system [34,35]. We employed the DMD device for simultaneous light splitting and spatial coding instead of traditional beam-splitter-based systems. The coding template is composed of equidistant horizontal fringes, with slightly larger spacing than the dispersion width, ensuring the maximum sampling rate and meeting the no-aliasing requirement. However, our RGB camera captures a different image from the HVIS system. Therefore, we utilized a bilateral filtering algorithm to fill the holes of the online prior channel and modify the algorithm for spectral data reconstruction [43]. This approach spreads the spectral information across the spatial and photometric domains, adapting to our system.

The captured patterns reflected from the DMD undergo spatial encoding, resulting in two sets of complementary patterns. Therefore, it is necessary to estimate the pixel values of the sampled points on the online prior channel using information from the sampled points on the dispersion channel to fill the missing sampled points in the RGB image. Figure 6 illustrates the correspondence between hyperspectral samples and RGB pixels, providing high-resolution spectral values for certain pixels in each RGB frame. Suppose we have an RGB frame with a resolution of $p \times q$ pixels and a corresponding hyperspectral frame of resolution $m \times n (m < p, n < q)$. Different from the HVIS system, where purple pixels (known SPE and RGB information) and blue pixels (unknown SPE but known RGB value) are known, our system has knowledge of blue pixels and green pixels (unknown RGB but known SPE value). For each blue pixel i, its hyperspectral information $I_{spe}(i)$ is determined by neighboring green pixels as

$$I_{spe}(i) = \frac{\sum_{j\in\Omega} D(i,j) \cdot I_{spe}(j)}{\sum_{j\in\Omega} D(i,j)}, \tag{1}$$

where $I_{spe}(i)$ denotes the hyperspectral vector for pixel i, $j \in \Omega$ indexes the green pixels within a neighborhood centered on j, and $D(i,j)$ represents the kernel function including the spatial domain and range domain as

$$D(i,j) = exp(-\frac{(i-j)^2}{2\sigma_d^2} - \frac{\parallel I_{RGB}(i) - F_j(I_{spe}(j)) \parallel^2}{2\sigma_r^2}), \tag{2}$$

where $I_{RGB}(i)$ denotes the RGB information of the blue pixel i, σ_d and σ_r are smoothing parameters, and $\parallel \cdot \parallel^2$ is the pixel-wise L_2 distance. $F_j(\cdot) \in R^{N\to 3}$ is a transfer function that estimates the 3-channel RGB pixel from N-dim spectral vector and can usually be obtained directly from the spectral response curve $T(\cdot)$ of the RGB camera. However, the diamond–square matching problem of the DMD and the detector makes it necessary to normalize $T(\cdot)$. Finally, $F_j(\cdot)$ can be expressed as the result of normalizing $T(\cdot)$ as

$$F_j(\cdot) = T(\cdot) \cdot \frac{\parallel I_{RGB}(j) \parallel}{\parallel I_{spe}(j) \parallel}, \tag{3}$$

where $I_{RGB}(j)$ denotes the pixel value obtained by transforming the spectral dimension vector from the response curve $T(\cdot)$; from that, we can calculate the RGB information regarding the green pixel j.

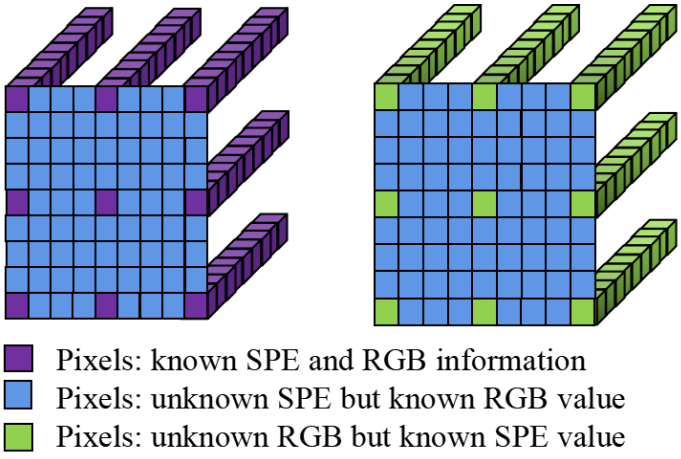


Figure 6. Schematic diagram of multispectral pixel distribution. The left picture is the reconstruction method adopted by the system using the beam splitter, and the RGB information of the sampling points in each spectral channel is known; the right picture is the DMD-CDM reconstruction method.

Filling the missing pixel values in the online prior channel is determined by the unique complementary and symmetrical characteristics of the DMD. The two-channel cameras jointly acquire all the necessary information through complementary sampling. To simplify this problem, we utilize the spectral channel to supplement the missing sampling points in the color channel, thereby transforming the task into a more general multispectral image fusion problem. The hyperspectral reconstruction problem is commonly formulated as the fusion of a low-resolution hyperspectral image (LR-HSI) and a high-resolution multispectral image (HR-MSI). To address the stringent real-time requirements of remote sensing applications, we adopted a lightweight multilevel information fusion network, specifically designed for hyperspectral and multispectral image fusion, as proposed in [48]. The architecture of the fusion network is depicted in Figure 7. Initially, the LR-HSI is upsampled using the bicubic interpolation method to match the spatial resolution of the

HR-MSI. The upsampled hyperspectral image (UP-HSI) is then concatenated with the HR-MSI to form the input for the fusion network.

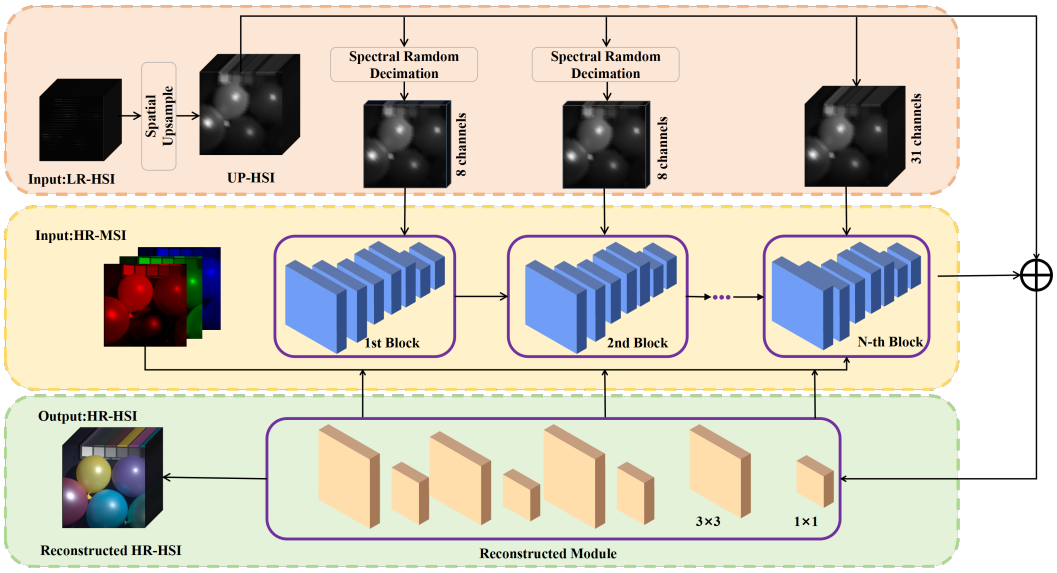


Figure 7. Schematic diagram of lightweight hyperspectral fusion network. UP-HSI is the upsampled image on LR-HSI, selecting spectral bands at equal intervals and sequentially performing feature extraction and fusion with HR-MSI image.

The network architecture is designed to effectively integrate spectral and spatial information for high-quality hyperspectral image (HSI) reconstruction. Specifically, the upsampled hyperspectral image (UP-HSI) is uniformly sampled at equal intervals and concatenated with the RGB image, which has been preprocessed to fill the missing points. The concatenated three-dimensional data cube is then fed into five residual feature extraction modules. Each module is composed of convolutional neural networks (CNNs) equipped with a channel attention mechanism, which enhances the extraction of discriminative spatial and spectral features.

During the fusion process, the UP-HSI is progressively fed into the feature extraction blocks at equal intervals, ensuring hierarchical refinement of the features. Each residual module is designed to process specific spectral bands and iteratively aggregate the outputs, maintaining consistency between the spatial and spectral domains throughout the fusion process. The output of the final residual module represents the fused spectral–spatial residual information. At this stage, the complete spectral bands of the UP-HSI are combined with the fusion residual via pointwise addition. The resulting data are then passed through a fine-tuning module composed of simple 1×1 and 3×3 convolutional layers. This module further optimizes the reconstruction process and generates the final high-resolution hyperspectral image (HR-HSI).

The integration of the channel attention mechanism within the CNN framework ensures that the network effectively captures both the spatial and spectral correlations, facilitating the fusion process. By leveraging this architecture, the proposed network achieves a balance between computational efficiency and reconstruction accuracy, producing high-quality HR-HSI while maintaining consistency across spatial and spectral domains.

3. Experiments and Analysis

In this section, we present the detailed evaluation of our proposed system, which integrates a DMD-based dual-channel imaging system with a lightweight multispectral fusion network. The experiments are divided into two parts: System Settings and Qualitative and Quantitative Evaluation. In the System Settings subsection, we describe the hardware configu-

ration, imaging parameters, and network implementation details to provide a comprehensive understanding of the experimental environment. In the Qualitative and Quantitative Evaluation subsection, we analyze the performance of the system through visual comparisons and numerical metrics, highlighting the effectiveness of the dual-channel imaging system and the designed fusion network in achieving high-quality multispectral reconstruction.

3.1. System Settings

Figure 1 presents a detailed schematic representation of our novel system, while Figure 8 showcases its physical implementation. Central to our design is the DMD, ingeniously serving dual roles as both a beam splitter and a spatial light modulator. We utilize a non-aliasing coding template within the DMD for scene selection and sampling; this template directs sampled light at a -12° angle towards the spatial coding channel and complementary sampling information at a $+12^\circ$ angle towards the online prior channel. In the spatial coding channel, a dispersive prism initially encodes the light before it is captured by the detector, facilitating the measurement of spectral details of the observed scene. Conversely, the online prior channel leverages color detectors to acquire complementary sampling information. This channel is designed to exhibit higher light throughput compared to its spatially coded counterpart, capturing more motion details in the measurement process. The convergence of data from both channels, underpinned by meticulous calibration and synchronization, enables the successful reconstruction of the underlying hyperspectral video.

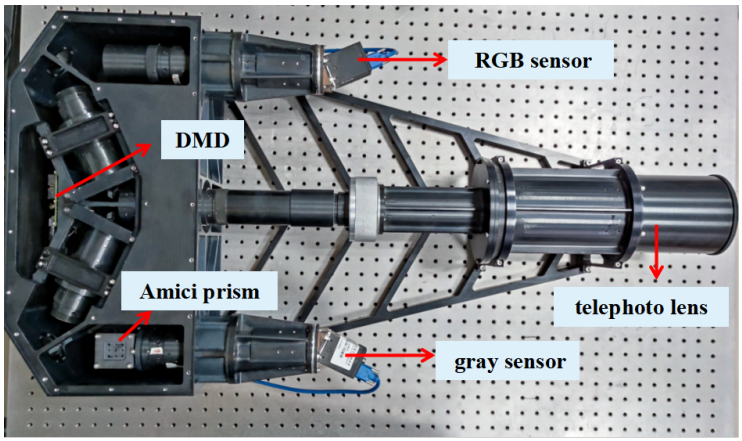


Figure 8. Physical diagram of the proposed system, consisting of an objective lens, a DMD, two sets of reflectors, a relay lens, an Amici prism, a grayscale camera, and an RGB camera.

In order to solve the component collision problem caused by the limited deflection angle of the DMD micromirror, we introduced two sets of reflectors in the spatial encoding channel and the online prior channel to change the optical path. The purpose of this design is to prevent the telecentric lens and filter in the imaging channel from colliding with the relay mirror in the reflection channel. Prior to operation, our system necessitates two critical procedures: calibration and synchronization. Systems based on beam splitters require individual calibration of each branch, followed by dual-channel registration. However, direct coordinate correspondence using feature point matching poses significant challenges due to non-parallel optical axes and inaccurate camera spacing. In our approach, we utilize the digital micromirror device (DMD) as a shared reflective surface for both optical paths in the computational spectral imaging system. This setup enables dual-channel pixel-level registration by mapping the coordinates between all pixels on the DMD and the two captured images. This innovative methodology streamlines the complex process of dual-channel registration, substantially enhancing the system's accuracy and efficiency. On

the other hand, the mechanical deflection of the DMD's micromirrors triggers the cameras in both channels, facilitating high-precision spectral encoding and imaging, and ensuring the accuracy and reproducibility of the data. Moreover, the rapid flipping characteristic of the DMD micromirrors enables swift template changes within a brief timeframe, leveraging the high-speed capabilities of the DMD.

Our system is adept at directly outputting hyperspectral images at video frame rates, a capability that we initially tested through hyperspectral imaging experiments. We employed the proposed system to capture hyperspectral videos of authentic scenes. A key feature of our setup is the use of a DMD device, which performs dual functions of spectroscopy and spatial coding, thereby ensuring high light efficiency. By leveraging the complementary encoding properties of DMD, DMD-CDH can achieve full-frame sampling. The fusion network trained on public datasets is used for inference in real-world scenarios.

Our experimental setup included two PointGray GS3-U3-41C6-NIR (Point Grey, Richmond, BC, Canada) cameras, each boasting a resolution of 2048 × 2048 pixels and a pixel pitch of 5.5 µm. The gray camera was designated for the spatial encoding channel, while the color camera was utilized for the online prior channel. Given that the spatial resolution hinges on the DMD's resolution of 912 × 1140, we successfully restored videos with 912 × 1140 × 45 bands (ranging from 450 nm to 900 nm) at a frame rate of 30 fps. This was achieved using a server equipped with an Intel Core i7-8700K CPU (Intel Corporation, Santa Clara, CA, USA) at 3.70 GHz, 32 GB RAM, and an NVIDIA GeForce RTX 2080Ti GPU (NVIDIA Corporation, Santa Clara, CA, USA).

The proposed lightweight multispectral fusion network is implemented using the PyTorch framework, with an ADAM optimizer [49], and undergoes 3000 training iterations with an initial learning rate of 10^{-3}. The network was trained and evaluated on the CAVE [50] and HARVARD [51] datasets. The HARVARD dataset contains 50 hyperspectral images (1024 × 1392 × 31, 420–720 nm), of which 30 are used for training and 20 are used for testing, while the CAVE dataset includes 32 hyperspectral images (512 × 512 × 31, 400–700 nm), of which 20 are used for training and 12 are used for testing. To supplement the missing RGB images, high-resolution multispectral images (HR-MSIs) were simulated by combining HSI with Nikon D700 camera (Nikon Corporation, Tokyo, Japan) response function [52], and low-resolution hyperspectral images (LR-HSIs) were generated by applying a 7 × 7 Gaussian blur kernel (mean 0; standard deviation 2) followed by spatial downsampling with stride. Experiments were conducted on NVIDIA GPU systems equipped with PyTorch 1.8.x and CUDA 11.1, and additional validation was performed using real images from our spectral imaging system to ensure robust and universal performance.

The experimental outcomes, illustrated in Figure 9, feature eight bands of the video with a 50 nm interval. We showcase two sets of hyperspectral reconstruction images. The brightness of the red scene at 650 nm–750 nm is significantly higher than that at 500 nm and 550 nm, verifying the feasibility of spectral resolution. The prism's dispersion width is set at 35 pixels, and our custom-designed coding template features a horizontal stripe pattern with a 40-pixel interval to mitigate aliasing effects. The DMD is directly linked to both cameras to facilitate shooting through a hard trigger, with both devices connected to the server. Unknown points in RGB images are supplemented in a dual-channel joint manner, utilizing a lightweight convolutional neural network for real-time spectral reconstruction. Notably, the integration of the beam splitter and spatial coding significantly simplifies calibration. We achieve pixel-level calibration between the spatial coding channel and the online prior channel by simply flipping the DMD's mirror.

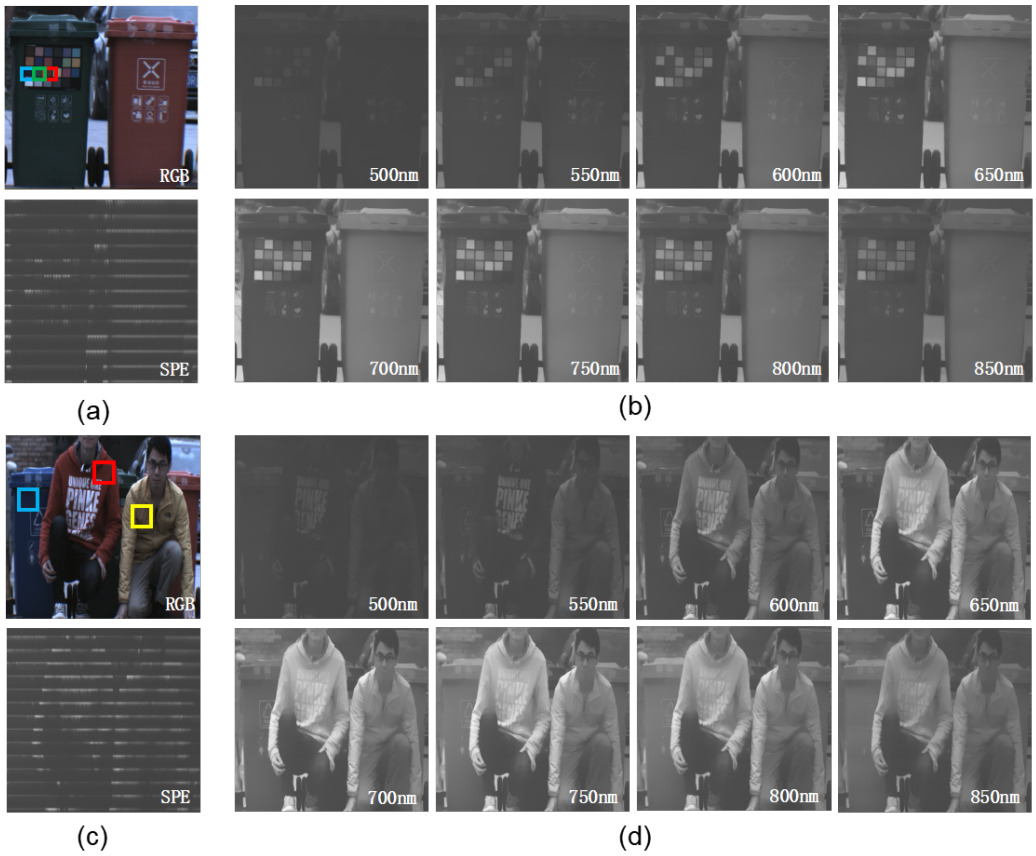
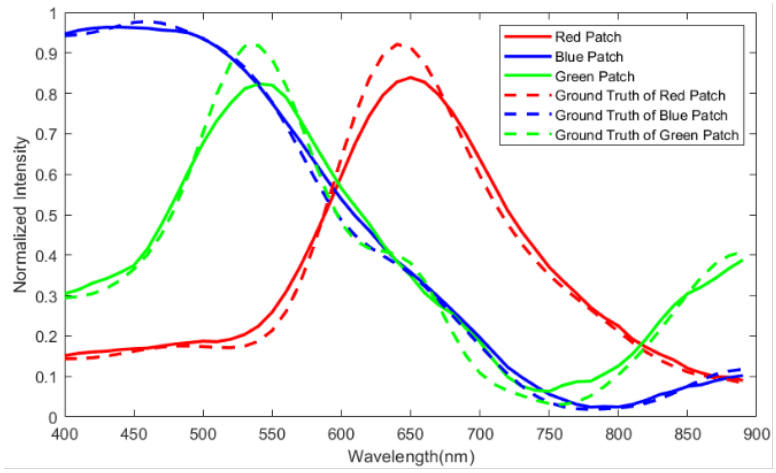


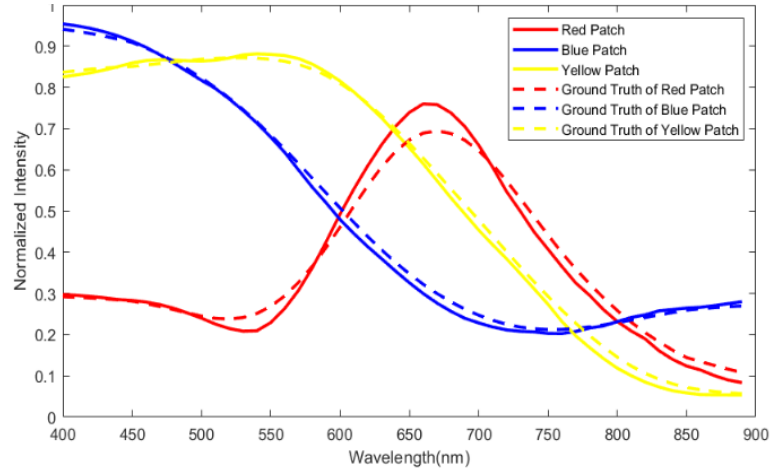
Figure 9. (**a**,**c**) Color image obtained after DMD encoding and reflection, grayscale image obtained after DMD-coded reflection, and prism dispersion. The different colored boxes in the figure correspond to the selected regions shown in Figure 10. (**b**,**d**) Reconstructed spectral image (wavelength from 500 to 850 nm; interval of 50 nm).

To verify the spectral accuracy of our imaging system, we used a commercial pushbroom spectrometer GaiaField to capture spectral information of the same target. GaiaField is a built-in pushscan hyperspectral camera with 128 channels and 696 × 700 spatial resolution. As shown in Figure 10, the comparison reveals a strong correlation between the average pixel information within the selected square and the ground truth. The dashed line represents the actual spectral response of the real scene, and the solid line is the system reconstructed spectrum response, demonstrating the reliable spectral accuracy of our system.

The spatial resolution of the sampled snapshot spectral imaging system is limited by the pixel size and resolution of the encoding template used in either the CASSI or HVIS systems. This is also true for our proposed hyperspectral imaging system, where the spatial resolution is the same as the DMD used, which is 912 × 1140. Using detectors with different resolutions does not improve the real spatial information. However, the spectral resolution of the spectral imaging system is determined by the dispersion width of the Amici prism and the calibration accuracy used. We used monochromatic light ranging from 450 to 900 nm to calibrate each by one, obtaining the coordinate mapping relationship under each wavelength, and were able to reconstruct 50 spectral bands in the spectral images.



(**a**)



(**b**)

Figure 10. Comparison curves between the spectral information of the three sample points intercepted and the ground truth. (**a**) Regions corresponding to the red, green, and blue boxes in Figure 9b; (**b**) Regions corresponding to the blue, red, and yellow boxes in Figure 9b.

3.2. Qualitative and Quantitative Evaluation

In this comparison, we evaluated our system against several computational imaging methods that are representative of the current hyperspectral imaging systems. The representative ones are CASSI and prism-mask spectral video imaging system (PMVIS) [35]. These systems can be broadly classified into two groups based on the coding templates used: CASSI-like and PMVIS-like. In CASSI-like systems, the objective lens first collects the spectral scene, which is then spatially coded using a fixed mask. The coded scene is spectrally dispersed using a prism or grating, and the resulting spatial–spectral coded scene is detected by a sensor. A snapshot of the sensor captures tens of spectral bands of the scene. In contrast, PMVIS-like systems directly acquire the spectra of scene points with the aid of a prism, using a mask with uniformly distributed holes to avoid overlaps of the dispersed spectra on the sensor. The spectral values of the sample points are directly measured, with a known correspondence between spectral bands and sensor plane locations. Both CASSI and PMVIS systems also use models of additional hybrid cameras. For the comparative test, we selected the single disperser CASSI [19], dual-CASSI [53], PMVIS [35], and HVIS [42] techniques. It is worth noting that both PMVIS and SD-CASSI systems place the mask on the imaging plane, resulting in pointwise coding.

We conducted a comparative analysis of five systems, focusing on differences in light throughput, the number of components used, the inclusion of prior information, the variability of encoding templates, Peak Signal-to-Noise Ratio (PSNR), and the time required to reconstruct a single frame. The accuracy of the reconstructed images is influenced by the varying spectral imaging systems and encoding methods, while the practical complexity of each system is largely dictated by the number of optical components and light flux.

Specifically, light throughput is a critical factor in computational hyperspectral imaging systems as it determines the overall percentage of light transmission through the system, primarily affected by light-blocking elements in the optical path. The comparison of light throughput across different systems, as shown in Figure 11, reveals that CASSI uses the fewest optical components, including a coding aperture with a 50 percent transmittance rate. For the PMVIS system, light throughput is governed by its subsampling rate, inversely proportional to the number of spectral channels. Dual-CASSI and HVIS employ beam splitters in conjunction with additional sensors to acquire prior information. However, the use of beam splitters results in a loss of half the light throughput. One channel utilizes an RGB camera, while the other shares the same light throughput with either CASSI or HVIS.

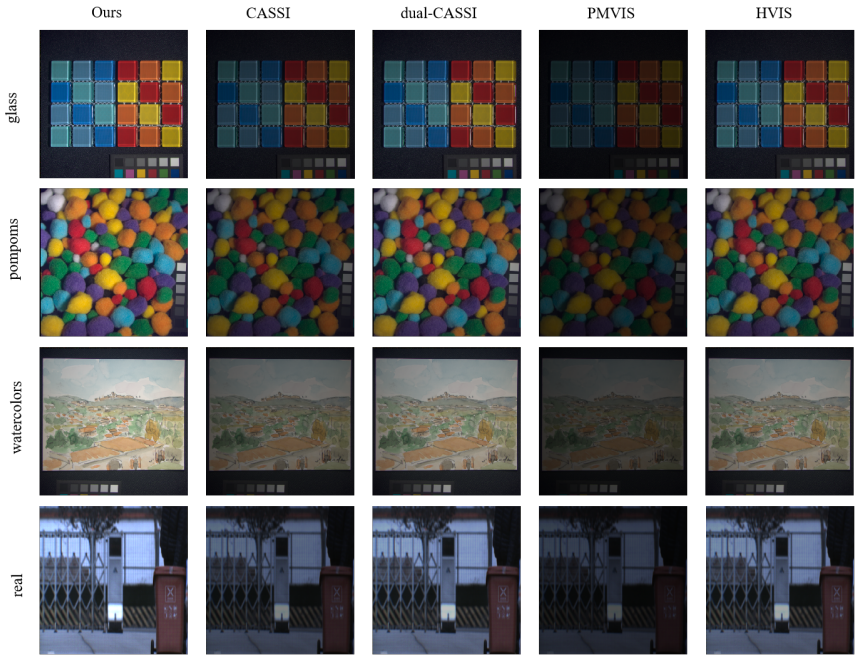


Figure 11. The comparison results of light throughput of different systems.

To thoroughly evaluate the novelty and superiority of the proposed hardware imaging system, we conducted a comparative study using the CAVE dataset and traditional reconstruction algorithms. Specifically, CASSI and dual-CASSI systems employed reconstruction algorithms grounded in compressive sensing theory, while PMVIS, HVIS, and our system utilized reconstruction methods based on unilateral or bilateral filtering techniques. The comparison results are presented in Table 2 and Figure 12, with the superior results highlighted for clarity.

The results presented in Table 2 demonstrate that our system consistently outperforms CASSI, primarily due to its use of a DMD-based coding device and its highly efficient light design. Notably, the reconstruction time of our system is approximately one-hundred-thousandth that required by CASSI, highlighting its significantly reduced computational complexity. While dual-CASSI achieves higher image quality compared to CASSI, its reliance on online a priori information from the panchromatic channel introduces increased computational complexity.

Table 2. Comparison of typical parameters with four types of undersampling systems. – indicates that this method does not have this functionality.

System	Light Throughput	Number of Devices	Prior Information	Change Template	PSNR (dB)	Time (s)
CASSI [19]	0.5	**6**	–	–	28.04	1173.385
dual-CASSI [53]	0.75	9	✓	–	32.10	1348.651
PMVIS [35]	$\frac{1}{\Omega}$	7	–	–	23.12	0.098
HVIS [42]	0.5 ($\frac{1}{\Omega}$ + 1)	9	✓	–	25.86	0.545
Ours	**1**	7	✓	✓	**32.19**	**0.064**

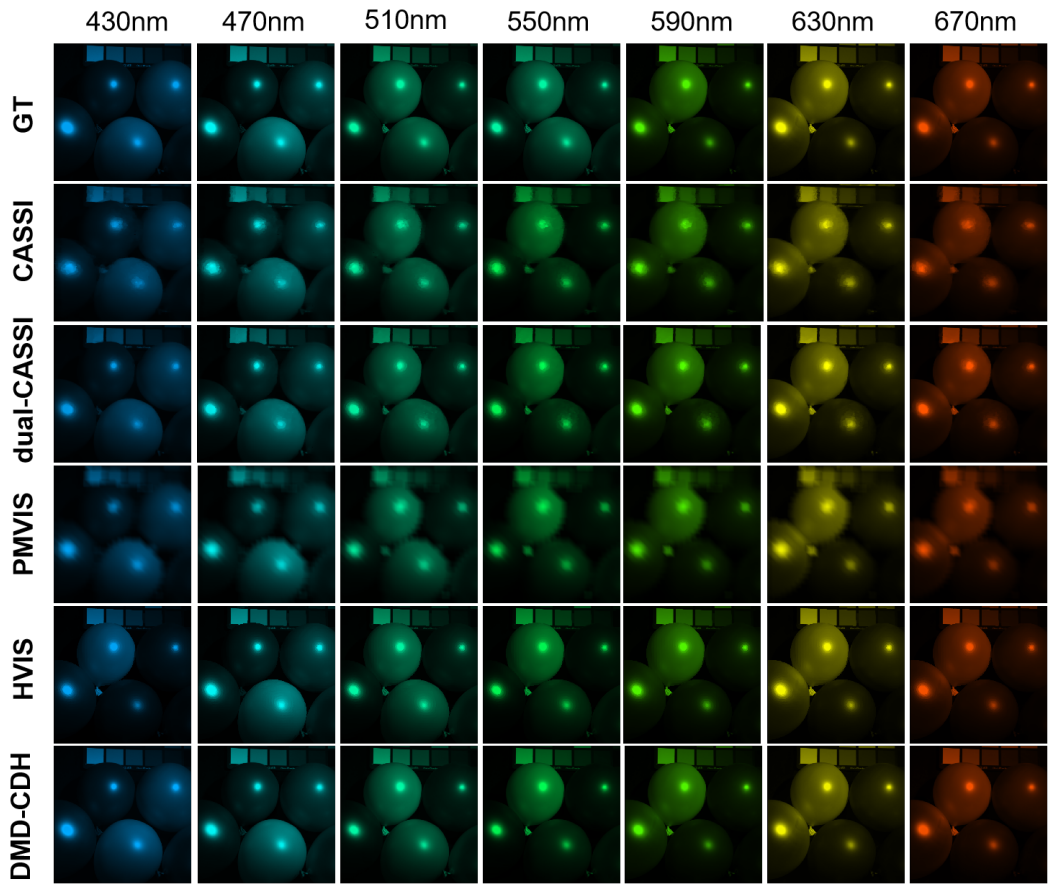


Figure 12. Difference results of "balloons-ms" image on the CAVE dataset at 430, 470, 510, 550, 590, 630, and 670 nm. The first row shows the ground-truth spectral band, and the other rows show the reconstructed results generated by the competing methods.

Both CASSI and HVIS systems incorporate a panchromatic camera to supplement spatial prior information during the imaging process. However, these systems are limited to capturing only half the incident light, which not only increases sensor noise but also reduces image smoothness. In contrast, our DMD-CDH system leverages DMD for both spectroscopic and spatial encoding, enabling the simultaneous capture of both panchromatic and spectral images of the same scene. Although dual-CASSI and HVIS demonstrate improvements in PSNR, the high complexity of dual-CASSI results in low optical throughput, large calibration errors, and offsets the potential advantages of its complex coded aperture system.

The proposed reconstruction algorithm further enhances performance by employing a convolutional neural network. Our system's innovative use of DMD enables the simultaneous capture of spatial scene information through complementary spatial coding and online prior channels. By doubling the incident energy and sampling rate, the system achieves

several advantages, including a smaller lens aperture, a higher frame rate, and simplified calibration through the integration of a beam splitter and spatial coding.

In practical implementations, pixel-level calibration between the spatial coding channel and the online prior channel is easily achieved by flipping the micromirror of the DMD. This simplifies the calibration process and enhances the system's versatility in real-world scenarios.

To verify the efficacy of the proposed method under smaller-sample circumstances, we conducted experiments using subsets of varying sizes from the CAVE datasets. The fusion network was fine-tuned with progressively reduced labeled training data, specifically using 100%, 75%, 50%, and 25% of the full dataset. The experimental results, summarized in Table 3, demonstrate that the system maintains robust performance even with significantly fewer training samples. Specifically, as the proportion of training data decreases, the PSNR remains above 48.99, the SSIM consistently exceeds 0.994, and the SAM stays within an acceptable range below 3.15. These results indicate a gradual decline in performance, but the system retains high reconstruction quality and structural similarity, underscoring its robustness. Notably, even with only 25% of the training data, the method achieves a PSNR of 48.99 and an SSIM of 0.9946, demonstrating its ability to operate effectively under constrained data conditions.

Table 3. Performance metrics of the proposed method on the CAVE dataset with varying training data proportions.

Training Data Proportion (%)	PSNR	SSIM	SAM	ERGAS
100%	50.2596	0.9958	2.8262	2.1676
75%	50.0992	0.9956	2.8421	2.1793
50%	49.8716	0.9954	2.9174	2.2346
25%	48.9935	0.9946	3.1481	2.3930

We summarize the comparison between the proposed method and three different types of imaging systems:

(1) Comparing with CASSI and its variants: DMD-CDH adopts the same Hadamard-like coding template in PMVIS to replace the complicated optimization compressive sensing restoration process and achieves real-time reconstruction.

(2) Comparing with PMVIS and its variants: PMVIS relies on the online prior channel to guide the super-resolution restoration. All the existing PMVIS-like systems use beam splitters to divide the incident light into two channels; thus, each channel can only obtain half the incident energy. DMD-CDH adopts the DMD device to accomplish both the beam splitting and spatial coding simultaneously, which can double the incident energy of the two channels.

(3) Comparing with other DMD-based computational imaging systems: There are two advantages of the proposed DMD-CDH system over the other current DMD-based systems. Firstly, the current DMD-based systems use a two-arm structure, one for imaging and the other for spatial coding. In other words, the DMD is only used as a changeable spatial coding template, and every micromirror is set to be ON or OFF according to the template. The light reflected by the OFF mirrors is wasted. DMD-CDH adopts a three-arm structure in which the reflected light of the OFF mirrors is recycled to provide an online prior channel. Secondly, the DMD-CDH system adopts the 'diamond-to-square-mirror-pixel' alignment setup, which eases the system assembling and commissioning process and avoids wasting either DMD micromirrors or sensor pixels.

4. Conclusions

In summary, we proposed a DMD-based complementary dual-channel hyperspectral video imaging system to address the key challenges of hyperspectral imaging, including high computational complexity, lack of online prior, low optical efficiency, and reliance on mechanical coding templates. Our approach reduces computational complexity by employing a lightweight hyperspectral fusion network, enabling real-time reconstruction at frame rates exceeding 30 fps. The dual-channel design leverages complementary reflections from the DMD to provide side-band information as an online prior, simplifying geometric calibration and enhancing reconstruction accuracy. By utilizing the high reflectivity and light efficiency of the DMD, our system achieves significant improvements in optical throughput compared to traditional systems. Furthermore, the reliance on a digital micromirror device eliminates the need for mechanical coding templates, ensuring stability and reliability in operation. These innovations not only improve the performance of hyperspectral video imaging but also lay the groundwork for future advancements in real-time high-performance hyperspectral systems. Future work will focus on improving imaging performance in extreme environments by utilizing dynamically adjustable coding templates and noise-adaptive algorithms, reducing alignment sensitivity through the introduction of larger-tilt-angle DMDs and improved calibration techniques, and optimizing neural networks to enhance accuracy and efficiency.

Author Contributions: Conceptualization, M.M., Y.N. and F.L.; methodology, D.G.; software, Y.N.; validation, F.L. and G.S.; formal analysis, M.M.; investigation, M.M. and Y.N.; resources, G.S.; data curation, F.L.; writing—original draft preparation, M.M. and D.G.; writing—review and editing, G.S.; visualization, M.M. and F.L.; supervision, Y.N. and G.S.; project administration, D.G.; funding acquisition, Y.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China under grants 62371356 and 62405238, in part by China Postdoctoral Science Foundation under grants 2024M752512, and the Key Research and Development Program of Shaanxi (No. 2023-YBNY-218 and 2023-YBNY-284).

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Fotiadou, K.; Tsagkatakis, G.; Tsakalides, P. Spectral super resolution of hyperspectral images via coupled dictionary learning. *IEEE Trans. Geosci. Remote Sens.* **2018**, *57*, 2777–2797. [CrossRef]
2. Du, X.; Zheng, X.; Lu, X.; Doudkin, A.A. Multisource Remote Sensing Data Classification with Graph Fusion Network. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 10062–10072. [CrossRef]
3. Gong, M.; Duan, Y.; Li, H. Group Self-Paced Learning With a Time-Varying Regularizer for Unsupervised Change Detection. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 2481–2493. [CrossRef]
4. Lian, J.; Wang, L.; Sun, H.; Huang, H. Gt-had: Gated transformer for hyperspectral anomaly detection. *IEEE Trans. Neural Netw. Learn. Syst.* **2024**, 1–15. [CrossRef]
5. Jalali, S.; Yuan, X. Snapshot compressed sensing: Performance bounds and algorithms . *IEEE Trans. Inf. Theory* **2019**, *65*, 8005–8024. [CrossRef]
6. Basedow, R.W.; Carmer, D.C.; Anderson, M.E.; Descour, M.R.; Mooney, J.M.; Perry, D.L.; Illing, L.R. HYDICE system: Implementation and performance. *Proc. SPIE-Int. Soc. Opt. Eng.* **1995**, *2480*, 258–267.
7. 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.* **2019**, *44*, 1281. [CrossRef]
8. Degraux, K.; Cambareri, V.; Geelen, B.; Jacques, L.; Lafruit, G. Multispectral Compressive Imaging Strategies Using Fabry–Pérot Filtered Sensors. *IEEE Trans. Comput. Imaging* **2018**, *4*, 661–673. [CrossRef]
9. Bian, L.; Wang, Z.; Zhang, Y.; Li, L.; Zhang, Y.; Yang, C.; Fang, W.; Zhao, J.; Zhu, C.; Meng, Q.; et al. A broadband hyperspectral image sensor with high spatio-temporal resolution. *Nature* **2024**, *635*, 73–81. [CrossRef] [PubMed]

10. Zhao, J.; Xu, H.; Liu, C.; Bian, L. Multimodal near-space surveillance system. In Proceedings of the Optoelectronic Imaging and Multimedia Technology XI, Nantong, China, 12–15 October 2024; SPIE: Bellingham, WA, USA, 2024; Volume 13239, pp. 251–258.
11. Wu, Z.; Lu, R.; Fu, Y.; Yuan, X. Latent Diffusion Prior Enhanced Deep Unfolding for Snapshot Spectral Compressive Imaging. In Proceedings of the European Conference on Computer Vision, Milan, Italy, 29 September–4 October 2024. [CrossRef]
12. Hong, D.; Wu, X.; Ghamisi, P.; Chanussot, J.; Yokoya, N.; Zhu, X.X. Invariant attribute profiles: A spatial-frequency joint feature extractor for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 3791–3808. [CrossRef]
13. Wang, L.; Zhang, T.; Fu, Y.; Huang, H. HyperReconNet: Joint Coded Aperture Optimization and Image Reconstruction for Compressive Hyperspectral Imaging. *IEEE Trans. Image Process.* **2019**, *28*, 2257–2270. [CrossRef] [PubMed]
14. Wang, H.; Wang, L.; Chen, C.; Hu, X.; Song, F.; Huang, H. Learning Spectral-wise Correlation for Spectral Super-Resolution: Where Similarity Meets Particularity. In Proceedings of the 31st ACM International Conference on Multimedia, Ottawa, ON, Canada, 29 October–3 November 2023; pp. 7676–7685.
15. Wang, Q.; Yuan, Z.; Du, Q.; Li, X. GETNET: A General End-to-End 2-D CNN Framework for Hyperspectral Image Change Detection. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 3–13. [CrossRef]
16. Wang, X.; Wang, L.; Ma, X.; Zhang, M.; Zhu, L.; Huang, H. In2SET: Intra-Inter Similarity Exploiting Transformer for Dual-Camera Compressive Hyperspectral Imaging. In Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Los Alamitos, CA, USA, 16–22 June 2024; pp. 24881–24891. [CrossRef]
17. Hagen, N.; Kudenov, M.W. Review of snapshot spectral imaging technologies. *Opt. Eng.* **2013**, *52*, 090901. [CrossRef]
18. Descour, M.; Dereniak, E. Computed-tomography imaging spectrometer: Experimental calibration and reconstruction results. *Appl. Opt.* **1995**, *34*, 4817–4826. [CrossRef]
19. Wagadarikar, A.; John, R.; Willett, R.; Brady, D. Single disperser design for coded aperture snapshot spectral imaging. *Appl. Opt.* **2008**, *47*, B44–B51. [CrossRef]
20. Lin, X.; Liu, Y.; Wu, J.; Dai, Q. Spatial-spectral encoded compressive hyperspectral imaging. *ACM Trans. Graph.* **2014**, *33*, 1–11. [CrossRef]
21. Ma, P.P.; Liu, Y.Y.; Lü, Q.; Pei, L.L.; Fang, Y. Optical System Design and Optimization of Compact Image Replicating Imaging Spectrometer. *Acta Photonica Sin.* **2016**, *45*. [CrossRef]
22. Qiao, M.; Liu, X.; Yuan, X. Snapshot spatial-temporal compressive imaging. *Opt. Lett.* **2020**, *45*, 1659–1662. [CrossRef] [PubMed]
23. Baek, S.H.; Kim, I.; Gutiérrez, D.; Min, H.K. Compact single-shot hyperspectral imaging using a prism. *ACM Trans. Graph. TOG* **2017**, *36*, 217. [CrossRef]
24. Li, W.; Hu, X.; Wu, J.; Fan, K.; Chen, B.; Zhang, C.; Hu, W.; Cao, X.; Jin, B.; Lu, Y.; et al. Dual-color terahertz spatial light modulator for single-pixel imaging. *Light Sci. Appl.* **2022**, *11*, 191. [CrossRef] [PubMed]
25. Jeon, D.S.; Baek, S.H.; Yi, S.; Fu, Q.; Min, H.K. Compact snapshot hyperspectral imaging with diffracted rotation. *ACM Trans. Graph.* **2019**, *38*, 1–13. [CrossRef]
26. Makarenko, M.; Burguete-Lopez, A.; Wang, Q.; Getman, F.; Giancola, S.; Ghanem, B.; Fratalocchi, A. Real-time hyperspectral imaging in hardware via trained metasurface encoders. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 12692–12702.
27. Liu, J.; Wu, Z.; Xiao, L.; Wu, X.J. Model inspired autoencoder for unsupervised hyperspectral image super-resolution. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–12. [CrossRef]
28. Chan, T.F.; Wong, C.-K. Total variation blind deconvolution. *IEEE Trans. Image Process.* **1998**, *7*, 370–375. [CrossRef]
29. Ahmed, N.; Natarajan, T.; Rao, K. Discrete Cosine Transform. *IEEE Trans. Comput.* **1974**, *C-23*, 90–93. [CrossRef]
30. Wang, L.; Sun, C.; Zhang, M.; Fu, Y.; Huang, H. Dnu: Deep non-local unrolling for computational spectral imaging. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 1661–1671.
31. Wang, Z.; Wang, L.; Huang, H. Sparse additive discriminant canonical correlation analysis for multiple features fusion. *Neurocomputing* **2021**, *463*, 185–197. [CrossRef]
32. Zheng, Y.; Li, J.; Li, Y.; Guo, J.; Wu, X.; Chanussot, J. Hyperspectral pansharpening using deep prior and dual attention residual network. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 8059–8076. [CrossRef]
33. Wang, Q.; Li, Q.; Li, X. A fast neighborhood grouping method for hyperspectral band selection. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 5028–5039. [CrossRef]
34. Hao, D.; Xin, T.; Xun, C.; Lin, S. A Prism-based System for Multispectral Video Acquisition. In Proceedings of the 2009 IEEE 12th International Conference on Computer Vision, Kyoto, Japan, 29 September–2 October 2009; pp. 175–182.
35. Cao, X.; Du, H.; Tong, X.; Dai, Q.; Lin, S. A Prism-Mask System for Multispectral Video Acquisition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *33*, 2423–2435. [CrossRef]
36. Boccalini, S.; Mayard, C.; Lacombe, H.; Villien, M.; Si-Mohamed, S.; Delahaye, F.; Boussel, L.; Budde, R.P.J.; Pozzi, M.; Douek, P. Ultra-High-Resolution and K-Edge Imaging of Prosthetic Heart Valves with Spectral Photon-Counting CT. *Investig. Radiol.* **2024**, *59*, 589–598. [CrossRef] [PubMed]

37. Dunlop-Gray, M.; Poon, P.K.; Golish, D.; Vera, E.; Gehm, M.E. Experimental demonstration of an adaptive architecture for direct spectral imaging classification. *Opt. Express* **2016**, *24*, 18307–18321. [CrossRef]
38. Lin, X.; Wetzstein, G.; Liu, Y.; Dai, Q. Dual-coded compressive hyperspectral imaging. *Opt. Lett.* **2014**, *39*, 2044. [CrossRef] [PubMed]
39. Wu, Y.; Mirza, I.O.; Arce, G.R.; Prather, D.W. Development of a digital micromirror device (DMD)-based Snapshot Spectral Imaging (DMD-SSI) system. In Proceedings of the IEEE Photonic Society 24th Annual Meeting Arlington, VA, USA, 9–13 October 2011; pp. 297–304.
40. Dong, X.; Xiao, X.; Pan, Y.; Wang, G.; Yu, Y. DMD-based hyperspectral imaging system with tunable spatial and spectral resolution. *Opt. Express* **2019**, *27*, 16995. [CrossRef] [PubMed]
41. Hellman, B.; Takashima, Y. Angular and spatial light modulation by single digital micromirror device for multi-image output and nearly-doubled étendue. *Opt. Express* **2019**, *27*, 21477–21496. [CrossRef]
42. Ma, C.; Cao, X.; Wu, R.; Dai, Q. Content-adaptive high-resolution hyperspectral video acquisition with a hybrid camera system . *Opt. Lett.* **2014**, *39*, 937–940. [CrossRef]
43. Cao, X.; Tong, X.; Dai, Q.; Lin, S. High resolution multispectral video capture with a hybrid camera system. In Proceedings of the Computer Vision Pattern Recognition, Colorado Springs, CO, USA, 20–25 June 2011; pp. 297–304.
44. Yang, S.; Yan, X.; Qin, H.; Zeng, Q.; Yuan, X. Mid-Infrared Compressive Hyperspectral Imaging. *Remote Sens.* **2021**, *13*, 741. [CrossRef]
45. Zhang, H.; Ma, X.; Lau, D.L.; Zhu, J.; Arce, G.R. Compressive Spectral Imaging Based on Hexagonal Blue Noise Coded Apertures. *IEEE Trans. Comput. Imaging* **2020**, *6*, 749–763. [CrossRef]
46. Chen, L.; Cai, L.; Huang, E.; Zhou, Y.; Yue, T.; Cao, X. A notch-mask and dual-prism system for snapshot spectral imaging. *Opt. Lasers Eng.* **2023**, *165*, 107544. [CrossRef]
47. Galvis, L.; Arguello, H.; Arce, G.R. Coded aperture design in mismatched compressive spectral imaging. *Appl. Opt.* **2015**, *54*, 9875. [CrossRef]
48. Ma, M.; Niu, Y.; Liu, C.; Li, F.; Shi, G. A Lightweight Multi-Level Information Network for Multispectral and Hyperspectral Image Fusion. *Remote Sens.* **2022**, *14*, 5600. [CrossRef]
49. Simoes, M.; Bioucas-Dias, J.; Almeida, L.B.; Chanussot, J. A Convex Formulation for Hyperspectral Image Superresolution via Subspace-Based Regularization. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 3373–3388. [CrossRef]
50. 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. Publ. IEEE Signal Process. Soc.* **2010**, *19*, 2241. [CrossRef] [PubMed]
51. Chakrabarti, A.; Zickler, T. Statistics of real-world hyperspectral images. In Proceedings of the IEEE Conference on Computer Vision Pattern Recognition, Colorado Springs, CO, USA, 20–25 June 2011; pp. 193–200
52. Wang, W.; Fu, X.; Zeng, W.; Sun, L.; Zhan, R.; Huang, Y.; Ding, X. Enhanced Deep Blind Hyperspectral Image Fusion. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *34*, 1513–1523. [CrossRef] [PubMed]
53. Wang, L.; Xiong, Z.; Gao, D.; Shi, G.; Zeng, W.; Wu, F. High-speed hyperspectral video acquisition with a dual-camera architecture. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 4942–4950.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.