# Mitigating Temporal Illumination Variation in Whisk-broom Hyperspectral Imaging using Dual-Perpendicular Scans

SUZAN JOSEPH KESSY[1,a)] TAKUYA FUNATOMI[1] KAZUYA KITANO[1]
YUKI FUJIMURA[1] TAKAHIRO KUSHIDA[1,2] GUILLAUME CARON[3,4]
EL MUSTAPHA MOUADDIB[3] YASUHIRO MUKAIGAWA[1]

## Abstract

This paper presents a method for temporal illumination variations compensation in whisk-broom hyperspectral imaging, targeting in-site captured images of stained-glass windows inside Amiens Cathedral, France. A previous method used a single vertical scan alongside the standard horizontal (raster) scan. However, it fails when the additional single-vertical scan is performed near or on a black frame. Our approach uses two scans scanned perpendicular to each other. We also challenge the assumption that illumination changes are negligible across columns and rows as proposed by the previous method; instead, we propose smooth illumination variations across columns and rows. Our method uses logarithmic space, leverages the low-dimensional structures of the logarithm of illumination and reflectance spectra, and enhances smoothness through total variation (TV) penalisation in the loss function. The quantitative evaluation of the compensation method using a spectral library and the field experiment results demonstrate our method's effectiveness in eliminating temporal illumination variations.

## 1. Introduction

In computer vision, the aim is to enhance computers' understanding of digital images or videos. This field has evolved from using traditional RGB (red, green, and blue) data, which captures information in three channels, to hyperspectral data that achieves superior spectral discrimination by capturing data across hundreds to potentially over 2000 narrow bands [7]. This makes hyperspectral data suitable for various applications, including agricultural monitoring [1], biomedical engineering [9], bio-metrics, and facial recognition [14] and cultural heritage analysis and preservation [2, 12].

Various methods exist for acquiring spectral images, including spectral reconstruction (SR) from RGB images [8], snapshot [5] and scanning techniques [11], each requiring a balance between temporal, spatial, and spectral resolutions. The whisk-broom scanning approach is favoured for its high spectral resolution and adaptable spatial scanning, despite its time-consuming process of scanning a scene pixel by pixel [7]. This method captures hyperspectral images with thousands of spectral channels but requires several hours to scan an entire scene [4]. Consequently, fluctuations in environmental illumination introduce temporal illumination variations affecting spectral cubes and hinder scene analysis.

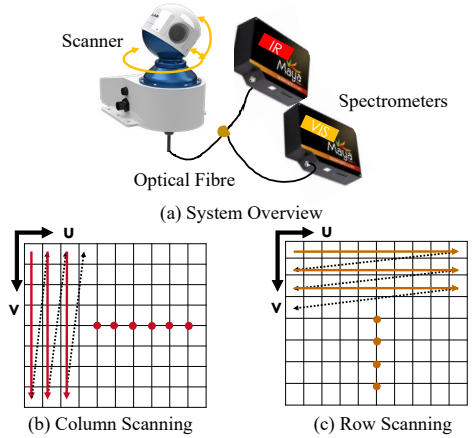


Fig. 1: (a) Overview of the imaging system designed for this study. (b) The whisk-broom column-wise scanning (c) The whisk-broom row-wise scanning to form a row-scan image.

This paper presents a method to compensate for changes in environmental illumination, building upon methods by [4, 7]. [4] addresses temporally varying illumination in horizontal (raster) scans by using a single column from a vertical scan. However, the method becomes more error-prone when the additional column scan is performed near or on a black frame. [7] proposed a solution using two perpendicular scans to tackle this issue yet this method introduced spatial artifacts in the compensated images. Both methods assumed negligible or constant illumination changes across rows or columns.

The present paper proposes a method of compensating for changes in environmental illumination using two perpendicular scans and unlike previous approaches, we assume the illumination variation as a smooth temporal gradient in each row or col-

1 Nara Institute of Science and Technology, Nara, Japan
2 Ritsumeikan University, Shiga, Japan
3 University of Picardie Jules Verne, MIS Laboratory, Amiens, France
4 CNRS-AIST (Joint Robotics Laboratory),IRL, Tsukuba, Japan
a) kessy.suzan_joseph.ks5@is.naist.jp

umn. We integrate a Total Variation (TV) regulariser in the loss function making the method robust against noise and prevents spatial artefacts, as validated by field and laboratory experiments. Key contributions include:

( 1 ) We introduce a method under the assumption that the illumination changes is smooth during each scan.

( 2 ) We integrated a TV regulariser to reduce noise and prevent artifacts during the compensation process.

( 3 ) We validated the method's effectiveness through field experiments on stained-glass windows in Amiens Cathedral, showcasing its superior performance.

## 2. Mitigating temporal illumination variation using dual perpendicular scans

### 2.1 Notations and problem setting

Let us consider a hyperspectral image $\mathcal{I} \in \mathbb{R}_+^{U\times V\times\Lambda}$, a 3D data cube captured by a hyperspectral camera, with a resolution of $U \times V \times \Lambda$, where $U$ and $V$ denote the width and height of the image, respectively and $\Lambda$ is the number of spectral channels. Let $\mathbf{i}_{u,v}$ denote a spectral vector at location $(u, v)$ we can decompose it as

$$\mathbf{i}_{u,v} = \boldsymbol{\rho}_{u,v} \odot \mathbf{l}, \tag{1}$$

where ($\boldsymbol{\rho}_{u,v} \in \mathbb{R}_+^{1\times\Lambda}$) is the *reflectance (or transmittance)* spectrum of the scene at point $(u, v)$, $\mathbf{l} \in \mathbb{R}_+^{1\times\Lambda}$ is the *illumination* spectrum, and $\odot$ represents the Hadamard (element-wise) product.

When there is fluctuation in the environmental illumination of a static scene, the illumination spectrum becomes a function of time denoted as $\mathbf{l}(t)$ and Eq. (1) can be rewritten as

$$\mathbf{i}_{u,v} = \boldsymbol{\rho}_{u,v} \odot \mathbf{l}(t), \tag{2}$$

where $t$ denotes the time taken for a scan. Our objective is to reconstruct the hyperspectral image, mitigating the impact of these variations.

### 2.2 Whisk-broom hyperspectral imaging

The devised whisk-broom imaging system contains a high spectral-resolution spectrometer with a two-dimensional mechanical scanning head as illustrated in Fig. 1(a). This spectrometer captures the spectral distribution of each discrete scene point, complemented by the flexible mechanical scanner head systematically sweeping the scene to construct a comprehensive hyperspectral image. The mechanical scanner operates a spatial scanning in the azimuth and elevation directions, thus the scanner can execute both horizontal (row-wise) and vertical (column-wise) scans of the scene. In the column-wise scanning mode, the sequential process advances from top to bottom, column by column, and from left to right, as shown by the red arrows in Fig. 1(b). Conversely, for row-wise, the progression moves from left to right, row by row, and from top to bottom, as shown by the orange arrows in Fig. 1(c).

### 2.3 Illumination reflectance spectra separation (IRSS)

We first separate the illumination and reflectance spectra and thereafter use the estimated values to mitigate temporal illumination variations.

#### 2.3.1 Assuming smooth transition of illumination changes in single line scan

Since the illumination $\mathbf{l}(t)$ and reflectance $\boldsymbol{\rho}_{u,v}$ spectra are unknown, the separation problem is non-unique. To simplify the problem, we assume that we can model the illumination variation in each single line (i.e. row or column) as a smooth temporal gradient.

Let $\mathbf{i}^{\rightarrow}$ and $\mathbf{i}^{\downarrow}$ denote row- and column-wise scans respectively. Given that the formulation of the hyper-spectral image follows a multiplicative model, as expressed in Eq. (2), we can transform the formulation into logarithm-spectral space and represent the first pixels of each scan in logarithmic space as

$$\begin{aligned} \log \mathbf{i}^{\rightarrow}_{1,v} &= \log \boldsymbol{\rho}_{1,v} + \log \mathbf{l}(t^{\rightarrow}_{1,v}) \\ \log \mathbf{i}^{\downarrow}_{u,1} &= \log \boldsymbol{\rho}_{u,1} + \log \mathbf{l}(t^{\downarrow}_{u,1}) \ , \end{aligned} \tag{3}$$

where $t$ is defined as

$$\begin{aligned} t^{\rightarrow}_{1,v} &\in \left[t^{\rightarrow}_{1,1} \cdots t^{\rightarrow}_{1,V}\right] \\ t^{\downarrow}_{u,1} &\in \left[t^{\downarrow}_{1,1} \cdots t^{\downarrow}_{U,1}\right]. \end{aligned} \tag{4}$$

We then linearly interpolate the illumination changes during the scan in the log-spectral space:

$$\begin{aligned} \log \tilde{\mathbf{l}}^{\rightarrow}_{u,v} &= \left(1 - w^{\rightarrow}_{u,v}\right) \log \mathbf{l}(t^{\rightarrow}_{1,v}) + w^{\rightarrow}_{u,v} \log \mathbf{l}(t^{\rightarrow}_{1,v+1}) \\ \log \tilde{\mathbf{l}}^{\downarrow}_{u,v} &= \left(1 - w^{\downarrow}_{u,v}\right) \log \mathbf{l}(t^{\downarrow}_{u,1}) + w^{\downarrow}_{u,v} \log \mathbf{l}(t^{\downarrow}_{u+1,1}), \end{aligned} \tag{5}$$

where $w^{\rightarrow}_{u,v} = \frac{t^{\rightarrow}_{u,v} - t^{\rightarrow}_{1,v}}{t^{\rightarrow}_{1,v+1} - t^{\rightarrow}_{1,v}}$, $w^{\downarrow}_{u,v} = \frac{t^{\downarrow}_{u,v} - t^{\downarrow}_{u,1}}{t^{\downarrow}_{u+1,1} - t^{\downarrow}_{u,1}}$. The scans are then written as

$$\begin{aligned} \log \mathbf{i}^{\rightarrow}_{u,v} &= \log \boldsymbol{\rho}_{u,v} + \log \tilde{\mathbf{l}}(t^{\rightarrow}_{u,v}) \\ \log \mathbf{i}^{\downarrow}_{u,v} &= \log \boldsymbol{\rho}_{u,v} + \log \tilde{\mathbf{l}}(t^{\downarrow}_{u,v}) \end{aligned} \ . \tag{6}$$

To interpolate all pixels, we introduced $\tilde{\mathbf{l}}(t^{\rightarrow}_{1,V+1})$ to the row-wise scan ($\mathbf{i}^{\rightarrow}_{u,v}$) and $\tilde{\mathbf{l}}(t^{\downarrow}_{U+1,1})$ for the column-wise scan ($\mathbf{i}^{\downarrow}_{u,v}$).

Building on the findings of [3,6], that showed illumination and reflectance spectra can be presented in low-dimensional spectral structure, we approximate $\log \boldsymbol{\rho}_{u,v}$ and $\log \mathbf{l}(t)$ as linear combinations of a limited set of basis spectra as

$$\log \boldsymbol{\rho}_{u,v} \simeq \boldsymbol{\psi}_{u,v}\mathbf{S}, \tag{7}$$

$$\log \mathbf{l}(t^{\rightarrow}_{1,v}) \simeq \boldsymbol{\xi}^{\rightarrow}_{v}\mathbf{E} \tag{8}$$

$$\log \mathbf{l}(t^{\downarrow}_{u,1}) \simeq \boldsymbol{\xi}^{\downarrow}_{u}\mathbf{E} \tag{9}$$

where $\boldsymbol{\psi}_{u,v} \in \mathbb{R}^{1\times M}, \boldsymbol{\xi}^{\rightarrow}_{v}, \boldsymbol{\xi}^{\downarrow}_{u} \in \mathbb{R}^{1\times N}$ are the weighting factor associated with basis spectrum $\mathbf{S} \in \mathbb{R}^{M\times\Lambda}$ and spectrum $\mathbf{E} \in \mathbb{R}^{N\times\Lambda}$ respectively. $M$, $N$ are the number of reflectance and illumination bases, respectively. We insert this into Eq. (6) and represent it in matrix form as

$$\mathbf{C} = \mathbf{AZB} = \left[\begin{array}{c|c|c} \mathbf{I}_{UV} & \mathbf{W}^{\rightarrow} & \mathbf{0} \\ \mathbf{I}_{UV} & \mathbf{0} & \mathbf{W}^{\downarrow} \end{array}\right] \begin{bmatrix} \boldsymbol{\Psi} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Xi}^{\rightarrow} \\ \mathbf{0} & \boldsymbol{\Xi}^{\downarrow} \end{bmatrix} \begin{bmatrix} \mathbf{S} \\ \mathbf{E} \end{bmatrix}, \tag{10}$$

where $\mathbf{C}$ is the vertical stack of $\log \mathbf{i}^{\rightarrow}_{u,v}$ and $\log \mathbf{i}^{\downarrow}_{u,v}$ forming $\mathbb{R}^{2UV\times\Lambda}$, $I_{UV}$ is an $UV \times UV$ identity matrix, $\mathbf{W}^{\rightarrow}$ and $\mathbf{W}^{\downarrow}$ are $UV \times (V + 1)$ and $UV \times (U + 1)$ matrices of

$$\begin{aligned} \mathbf{W}^{\rightarrow}\left[u + vU,\, v'\right] &= \left(1 - w^{\rightarrow}_{u,v}\right)\delta_{v,\,v'} + w^{\rightarrow}_{u,v}\delta_{v+1,\,v'} \\ \mathbf{W}^{\downarrow}\left[u + vU,\, u'\right] &= \left(1 - w^{\rightarrow}_{u,v}\right)\delta_{u,\,u'} + w^{\rightarrow}_{u,v}\delta_{u+1,\,u'} \end{aligned}$$

$\delta_{i,j}$ is the Kronecker delta and $\boldsymbol{\Psi}, \boldsymbol{\Xi}^{\rightarrow}, \boldsymbol{\Xi}^{\downarrow}$ are the vertical stacks of $\boldsymbol{\psi}_{u,v}, \boldsymbol{\xi}_{v}^{\rightarrow}, \boldsymbol{\xi}_{u}^{\downarrow}$, respectively. Thus, $\boldsymbol{\Psi} \in \mathbb{R}^{UV \times M}$, $\boldsymbol{\Xi}^{\rightarrow} \in \mathbb{R}^{(V+1) \times N}$ and $\boldsymbol{\Xi}^{\downarrow} \in \mathbb{R}^{(U+1) \times N}$.

To estimate $\boldsymbol{\rho}_{u,v}$ and $\tilde{\mathbf{l}}(t)$, we need to solve Eq. 10 for $\mathbf{Z}$ and $\mathbf{B}$. By leveraging on illumination and reflectance spectra low-rank characteristics, we apply Singular Value Decomposition (SVD) to the observation matrix $\mathbf{C}$. This process factorizes $\mathbf{C}$ into

$$\mathbf{C} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\top} = \sum_{r=1}^{R} \mathbf{u}_r \sigma_r \mathbf{v}_r^{\top}, \tag{11}$$

where $\mathbf{U}$ and $\mathbf{V}$ are the orthogonal matrices, and $\boldsymbol{\Sigma}$ is the diagonal matrix containing the singular values. $\mathbf{u}_r$ and $\mathbf{v}_r$ are the $r$-th columns of $\mathbf{U}$ and $\mathbf{V}$, respectively, $\sigma_r$ is the $r$-th singular value, and $R$ is the rank of $\mathbf{C}$.

Note that $\mathbf{C}$ can be represented by Eq. (10) and Eq. (11) we can thus equate these two equations as $\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\top} = \mathbf{A}\mathbf{Z}\mathbf{B}$. In practice, basis spectrum $\mathbf{S}$ and $\mathbf{E}$ are not independent. Thus, $\mathbf{B} \neq \mathbf{V}^{\top}$. Now, let $\mathbf{B}'$ be a compressed basis for both illumination and reflectance spectra. By using the mixed weight matrix $\mathbf{Z}'$, the following equation can hold true: $\mathbf{C} = \mathbf{A}\mathbf{Z}'\mathbf{B}'$. Now, SVD can provide the bases as

$$\begin{aligned} \mathbf{A}\mathbf{Z}' &= \mathbf{U}\boldsymbol{\Sigma}, \\ \mathbf{B}' &= \mathbf{V}^{\top}. \end{aligned} \tag{12}$$

Since $\mathbf{A}$ is large and sparse, making its inverse calculation unstable, we directly solve $\mathbf{Z}'$ iteratively by minimizing $\|\mathbf{A}\mathbf{Z}^{'} - \mathbf{U}\boldsymbol{\Sigma}\|$.

From Eqs. (7), (8), and (9), we can estimate the logarithmic spectra for illumination and reflectance by

$$\begin{bmatrix} \mathbf{P} \\ \mathbf{G}^{\rightarrow} \\ \mathbf{G}^{\downarrow} \end{bmatrix} = \mathbf{Z}'\mathbf{V}^{\top}. \tag{13}$$

where $\mathbf{P}, \mathbf{G}^{\rightarrow}, \mathbf{G}^{\downarrow}$ are the vertical stack of $\log \boldsymbol{\rho}_{u,v}$, $\log \bar{\mathbf{l}}(t_{1,v}^{\rightarrow})$, $\log \bar{\mathbf{l}}(t_{u,1}^{\downarrow})$, respectively.

#### 2.3.2 Intensity filtering for the process in logarithmic space

Using logarithmic space requires careful handling of negative or small values for precision, as highlighted by [3]. Rapid sunlight variations can cause pixel saturation and clipping intensity at maximum values. To address this, we use intensity filtering, as outlined by [7]. After filtering, recovering reflectance from filtered pixels becomes impossible, making the estimated reflectance in Eq. (13) incomplete. However, the estimated illumination spectrum covers the entire row and column array since not all pixels are filtered in each row or column. Thus, we can infer the complete scene's reflectance using the estimated illumination spectra and the row-wise scan.

#### 2.3.3 Total variation loss function

To estimate the complete reflectance of the scene we incorporate a TV regulariser in the cost function to reinforce the reliability of the estimation procedure. The TV model's basis is that the integral of a signal's absolute gradient is high when noisy. Minimizing this integral reduces noise and smoothens adjacent pixels [13]. Let $H^{\rightarrow}$ denote the TV linear operator calculating the differences between the vectors corresponding to adjacent pixels


Fig. 2: Experimental setup: The object is a colour chart of ColorChecker Passport (X-rite, Inc) under uncontrollable illumination near a window.

in horizontal direction, and $H^{\downarrow}$ computes the vertical difference defined as

$$\begin{cases} H^{\rightarrow}(\mathbf{P}) = \sum_{u,v} \|\boldsymbol{\rho}_{u+1,v} - \boldsymbol{\rho}_{u,v}\|_1 \\ H^{\downarrow}(\mathbf{P}) = \sum_{u,v} \|\boldsymbol{\rho}_{u,v+1} - \boldsymbol{\rho}_{u,v}\|_1 \\ H^{\downarrow}(\mathbf{G}^{\rightarrow}) = \sum_{u,v} \|\mathbf{l}_{u,v+1} - \mathbf{l}_{u,v}\|_1 , \end{cases} \tag{14}$$

where $\|\cdot\|_1$ is the $\ell_1$ norm. The overall minimisation is given as

$$\begin{aligned} \hat{\mathbf{P}}, \hat{\mathbf{G}}^{\rightarrow} = \arg\min_{\{\boldsymbol{\rho}, \mathbf{l}\}} \Bigg( & \frac{1}{UV} \sum_{u,v} \left\| \mathbf{i}_{u,v}^{\rightarrow} - \boldsymbol{\rho}_{u,v} \odot \mathbf{l}_{u,v} \right\|_2^2 \\ & + \alpha \left( H^{\rightarrow}(\mathbf{P}) + H^{\downarrow}(\mathbf{P}) \right) + \beta H^{\downarrow}(\mathbf{G}^{\rightarrow}) \\ & + \nu \frac{1}{UV} \sum_{u,v} \left\| \mathbf{l}_{u,v} - \hat{\mathbf{l}}(t_{u,v}^{\rightarrow}) \right\| \Bigg), \end{aligned} \tag{15}$$

where $\alpha, \beta$ and $\nu$ are constants. We use Adam optimizer to achieve this minimization.

### 2.4 Mitigating temporal illumination variation

Our main goal is to achieve uniform illumination in the image. To compensate for illumination variations, we scale the estimated reflectance $\hat{\boldsymbol{\rho}}_{u,v}$ with a single row of the estimated illumination $\hat{\mathbf{l}}(t_{1,1}^{\rightarrow})$. The compensated vector $\hat{\mathbf{i}}$ is expressed as

$$\hat{\mathbf{i}}_{u,v} = e^{\hat{\boldsymbol{\rho}}_{u,v} \odot \hat{\mathbf{l}}(t_{1,1}^{\rightarrow})}. \tag{16}$$

## 3. Experiments

A series of experiments were carried out to evaluate the effectiveness and performance of the proposed method comparing it with the Robust Principal Component Analysis (RPCA) proposed by [4] and the Constant Illumination Variation (CIV) method proposed by [7].

### 3.1 Quantitative Evaluation

A custom imaging system with a scanning head and two

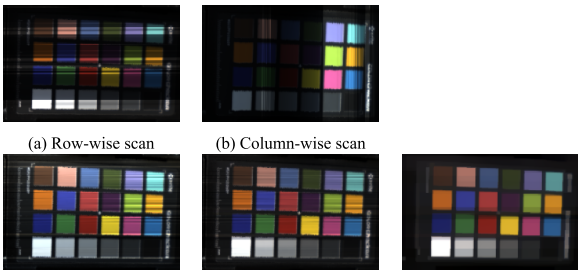


Fig. 3: RGB images were synthesised from scanned images with illumination variations and mitigated results.

Maya2000 Pro spectrometers (Ocean Optics, Inc.) was developed as shown in Fig. 1 (a). Spectrometers have 2068 channels, covering $199.50\,\mathrm{nm}$ to $1118.15\,\mathrm{nm}$ for visible range (*"VIS"*) and $500.81\,\mathrm{nm}$ to $1398.87\,\mathrm{nm}$ for Infrared (*"IR"*) range, with a $0.5\,\mathrm{nm}$ spectral resolution. The RobotEye REHS25 scanner (Ocular Robotics Ltd) achieves up to $0.01^\circ$ spatial resolution over $[-180^\circ, 180^\circ]$ in azimuth and $[-35^\circ, 35^\circ]$ in elevation directions. The optical connectivity was established by an optical fiber, with a $200\,\mu\mathrm{m}$ core.

Two hyperspectral images of a ColorChecker Passport were acquired under partially cloudy conditions, exhibiting varying sunlight as illustrated by Fig. 2. A row- scan and a column-wise scan were captured, with scan duration of $1197\,\mathrm{s}$ and $1275\,\mathrm{s}$, respectively. We used $50$ as the lower limit of the *intensity filter* and the upper limit was set as $5000$ for which $1351/4136$ channels were filtered out, $\alpha = 2\times10^{-8}$, $\beta = 6\times10^{-8}$, and $\nu = \beta = 1\times10^{2}$. To implement the RPCA, we used column $U = 68$, avoiding any near or black columns. Fig. 3 shows RGB images synthesised from spectral cubes using an open-source colour science package [10], comparing row- and column-wise scans, RPCA [4], CIV [7] and the proposed method. We can see temporal illumination variations are evident under natural sunlight on the scanned images and are effectively mitigated by the proposed method, outperforming RPCA [4] and CIV [7] in avoiding spatial artefacts.

#### 3.1.1 Evaluation using spectral library

The goal here is to compare spectra from color patches in spectral cubes with their counterparts in the Chromaxion spectral library, which contains reflectance values for ColorChecker Passport patches (X-Rite, Inc.) as introduced by [10]. We applied the pre-processing method from [4] for comparison and evaluated the results using Root Mean Squared Error (RMSE), calculating error for each pixel within the patches against the spectral library values across all channels. Fig. 4 (a–e) show error distributions across the cubes with mean RMSE indicated below each image. Row and column-wise scans exhibit significant errors due to illumination variations, while compensated results show reduced errors. These results demonstrate that employing dual perpendicular scans minimizes errors more effectively than a single vertical scan. Fig. 4 (f) compares error rates between CIV [7] and the proposed method (smooth varying), with blue highlighting regions where the proposed method outperforms CIV [7].

### 3.2 Experiments with Stained glass windows

We captured sunlight transmitted through stained glass windows in the *Cathédrale Notre-Dame d'Amiens*, located in Amiens, France. Both row- and column-wise scan images of stained glass window number **XVIII** were taken under varying temporal sunlight conditions. We applied mitigation methods to the hyperspectral images. For the filter, we set lower and upper limits at $0.5$ and $1000$, respectively, filtering out $1175/4136$ channels. Due to a lack of ground-truth hyperspectral images, only qualitative assessments were possible. Fig. 5 (a–c) show synthesised RGB images and mitigated results. The difference between the scan and the proposed method's result is visualised in Fig. 5 (d). Compared to the scan, the mitigated result exhibits lower intensity in blue and higher intensity in red areas.

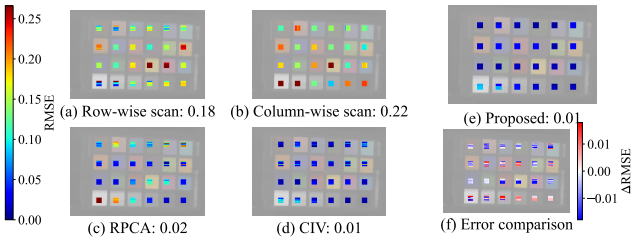


Fig. 4: The error maps of (a) the row-wise and (b) column-wise scans and the mitigated results from (c) the RPCA (d) the Constant illumination variation and (e) the proposed method. (f) Comparison of the errors from the mitigation methods.


Fig. 5: (a–c) Synthesized RGB image of the stained glass window number **XVIII** (d) the difference between the row-wise scan and the mitigated results.

## 4. Conclusion

This paper introduces a novel method combining whisk-broom imaging with total variation (TV) regularization to create a hyperspectral cube adaptable to varying illumination conditions. We employ two full scene scans, performed perpendicular to each other, demonstrating superiority over a single vertical scan (RPCA) method. Additionally, assuming smooth illumination variation yields RMSE errors comparable to those assumed under constant illumination variation. However, our method does not exhibit spatial artifacts.

## Acknowledgement

This work was partly supported by JSPS KAKENHI JP23H00499 and JST PRESTO JPMJPR2025.

## References


[1] Adam, E., Mutanga, O. and Rugege, D.: Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review, *Wetlands Ecology and Management*, Vol. 18, No. 3, pp. 281–296 (2010).

[2] Cutajar, J. D., Babini, A., Deborah, H., Hardeberg, J. Y., Joseph, E. and Frøysaker, T.: Hyperspectral Imaging Analyses of Cleaning Tests on Edvard Munch's Monumental Aula Paintings, *Studies in Conservation*, Vol. 67, No. sup1, pp. 59–68 (2022).

[3] Drew, M. S. and Finlayson, G. D.: Analytic solution for separating spectra into illumination and surface reflectance components, Vol. 24, No. 2, pp. 294–303 (2007).

[4] Funatomi, T., Ogawa, T., Tanaka, K., Kubo, H., Caron, G., Mouaddib, E. M., Matsushita, Y. and Mukaigawa, Y.: Eliminating Temporal Illumination Variations in Whisk-broom Hyperspectral Imaging, *International Journal of Computer Vision*, Vol. 130, No. 5, pp. 1310–1324 (2022).

[5] Hagen, N. A. and Kudenov, M. W.: Review of snapshot spectral imaging technologies, *Optical Engineering*, Vol. 52, No. 9, p. 090901 (2013).

[6] Judd, D. B., MacAdam, D. L., Wyszecki, G., Budde, H. W., Condit, H. R., Henderson, S. T. and Simonds, J. L.: Spectral Distribution of Typical Daylight as a Function of Correlated Color Temperature, Vol. 54, No. 8, pp. 1031–1040 (1964).

[7] Kessy, S. J., Funatomi, T., Kitano, K., Fujimura, Y., Caron, G., Mouaddib, E. M. and Mukaigawa, Y.: Hyperspectral Imaging of In-Site Stained Glasses: Illumination Variation Compensation Using Two Perpendicular Scans, *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, pp. 1662–1670 (2023).

[8] Li, J., Leng, Y., Song, R., Liu, W., Li, Y. and Du, Q.: MFormer: Taming Masked Transformer for Unsupervised Spectral Reconstruction, *IEEE Transactions on Geoscience and Remote Sensing* (2023).

[9] Lu, G. and Fei, B.: Medical hyperspectral imaging: a review, *Journal of Biomedical Optics*, Vol. 19, No. 1, p. 010901 (2014).

[10] Mansencal, T., Mauderer, M., Parsons, M., Shaw, N., Wheatley, K., Cooper, S., Vandenberg, J. D., Canavan, L., Crowson, K., Lev, O., Leinweber, K., Sharma, S., Sobotka, T. J., Moritz, D., Pppp, M., Rane, C., Eswaramoorthy, P., Mertic, J., Pearlstine, B., Leonhardt, M., Niemitalo, O., Szymanski, M., Schambach, M., Huang, S., Wei, M., Joywardhan, N., Wagih, O., Redman, P., Goldstone, J. and Hill, S.: Colour 0.3.16 (2020).

[11] Min, D., Zhao, J., Bodner, G., Ali, M., Li, F., Zhang, X. and Rewald, B.: Early decay detection in fruit by hyperspectral imaging–Principles and application potential, *Food Control*, Vol. 152, p. 109830 (2023).

[12] Pillay, R., Hardeberg, J. Y. and George, S.: Hyperspectral imaging of art: Acquisition and calibration workflows, *Journal of the American Institute for Conservation*, Vol. 58, No. 1-2, pp. 3–15 (2019).

[13] Rudin, L. I., Osher, S. and Fatemi, E.: Nonlinear total variation based noise removal algorithms, *Physica D: Nonlinear Phenomena*, Vol. 60, No. 1, pp. 259–268 (1992).

[14] Xie, Z., Li, Y., Niu, J., Shi, L., Wang, Z. and Lu, G.: Hyperspectral face recognition based on sparse spectral attention deep neural networks, *Optics Express*, Vol. 28, No. 24, pp. 36286–36303 (2020).