Correction of Dropped Frames in High-resolution Push-broom Hyperspectral Images for Cultural Heritage

CHARLIE WILLARD, University College London
NANCY WADE, Guildhall Art Gallery, City of London
MATIJA STRLIČ, University College London
JOHN R. GILCHRIST, Clyde Hyperspectral Imaging and Technology
TIM WEYRICH and ADAM GIBSON, University College London

Dropped frames can occur in line-scan cameras, which result in non-uniform spatial sampling of the scene. A dropped frame occurs when data from an image sensor is not successfully recorded. When mosaicking multiple line-scan images, such as in high-resolution imaging, this can cause misalignment. Much previous work to identify dropped frames in video prioritises fast computation over high accuracy, whereas in heritage imaging, high accuracy is often preferred over short computation time. Two approaches to identify the position of dropped frames are presented, both using the A* search algorithm to correct dropped frames. One method aligns overlapping sections of push-broom images and the other aligns the push-broom image to a lower resolution reference image. The two methods are compared across a range of test images, and the method aligning overlapping sections is shown to perform better than the method using a reference image under most circumstances. The overlap method was applied to hyperspectral images acquired of *La Ghirlandata,* an 1873 oil on canvas painting by D. G. Rossetti, enabling a high-resolution hyperspectral image mosaic to be produced. The resulting composite image is 10,875 × 14,697 pixels each with 500 spectral bands from 400–2,500 nm. This corresponds to a spatial resolution of 80 μm and a spectral resolution of 3–6 nm.

CCS Concepts: • **Applied computing** → **Fine arts**; • **Computing methodologies** → **Computational photography**; **Image processing;**

Additional Key Words and Phrases: Hyperspectral image, push-broom, line-scan, dropped frames, A* search, image stitching

ACM Reference format:
Charlie Willard, Nancy Wade, Matija Strlič, John R. Gilchrist, Tim Weyrich, and Adam Gibson. 2022. Correction of Dropped Frames in High-resolution Push-broom Hyperspectral Images for Cultural Heritage. *J. Comput. Cult. Herit.* 15, 2, Article 29 (April 2022), 19 pages.
https://doi.org/10.1145/3479011

1 INTRODUCTION

Hyperspectral imaging acquires information at high spectral resolution for every pixel in an image. There are multiple approaches to acquiring hyperspectral images, including spectral scanning [Gat 2000], spatio-spectral

Authors' addresses: C. Willard, University College London, Gower Street, London, WC1E 6BT; email: Charles.Willard.17@ucl.ac.uk; N. Wade, Guildhall Art Gallery, Guildhall Yard, London, EC2V 5AE; email: Nancy.Wade@cityoflondon.gov.uk; M. Strlic, University College London, Gower Street, London, WC1E 6BT; email: M.Strlic@ucl.ac.uk; J. R. Gilchrist, Clyde Hyperspectral Imaging and Technology, 1 Aurora Avenue, Clydebank, G81 1BF; email: Johnrg@clydehsi.com; T. Weyrich, University College London, Gower Street, London, WC1E 6BT; email: t.weyrich@cs.ucl.ac.uk; A. Gibson, University College London, Gower Street, London, WC1E 6BT; email: Adam.Gibson@ucl.ac.uk.

1556-4673/2022/04-ART29 $15.00
https://doi.org/10.1145/3479011

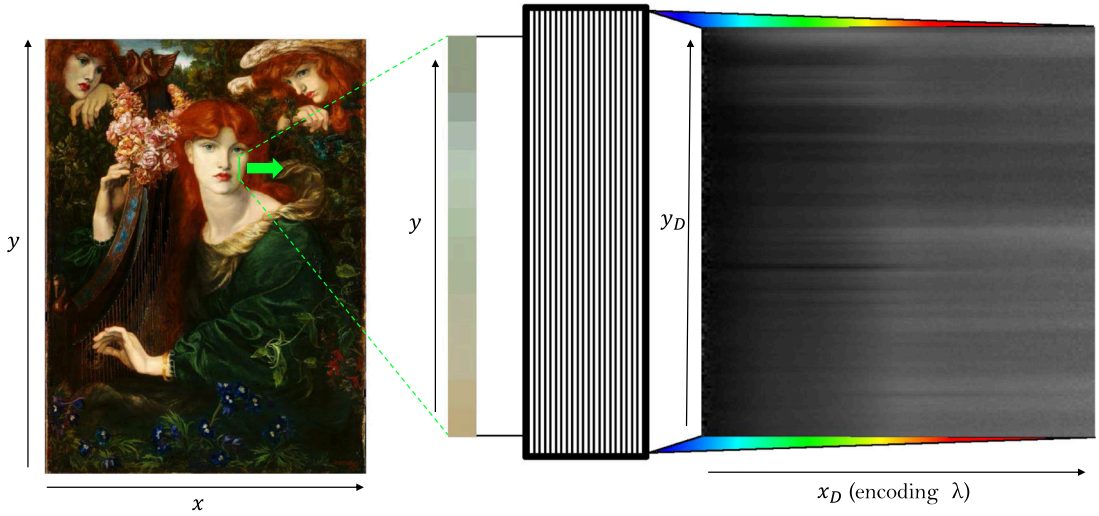


Fig. 1. A diagram showing push-broom hyperspectral image acquisition. We seek to acquire a data cube with three dimensions (x, y, λ). The image detector only has two spatial dimensions (x_D, y_D). To record the three-dimensional data cube, we identify a single one-dimensional line of pixels along the y direction and split this into its spectral components using a diffraction grating. The image sensor then encodes y along the detector's spatial dimension y_D and λ along the detector's spatial dimension x_D. As we scan in time, the slit moves across the painting and the second spatial dimension x is encoded in time t. In this way $y \rightarrow y_D, x \rightarrow t, \lambda \rightarrow x_D$. A single frame is defined as one exposure of the image detector (x_D, y_D) recording spatial information y and spectral information λ at one point along the spatial dimension x.

scanning [Renhorn et al. 2016], snapshot imaging [Bodkin et al. 2009], and spatial scanning [Yang et al. 2003]. High-resolution hyperspectral imaging of paintings is an active area of research, with line-scanning hyperspectral systems being used to collect large datasets [Favero et al. 2017; Gabrieli 2021]. Line-scanning systems use an imaging spectrograph to record spectral information across one axis and spatial information across the second axis of an imaging sensor. Line-scan hyperspectral cameras operate as spatial scanning systems and acquire one line of pixels at a time, building up an image as the camera moves across the scene, as shown in Figure 1. Each line of pixels is acquired from one exposure, referred to as a "frame," of the image sensor. A single "frame," therefore, is an image where a line of image pixels is represented along one spatial dimension, the wavelength of the pixels is represented along the second spatial dimension, and the spectral intensity is recorded by the intensity of each pixel. The full "image cube" is assembled from many frames and has three dimensions: two spatial dimensions and one spectral dimension, which can be calibrated to give spectral reflectance measurements. Data is transferred from the sensor often at a rate of tens of frames per second. In this way, a hyperspectral image cube is equivalent to a video where time corresponds to the camera movement across the scene. Dropped frames are known to occur in videos, particularly in video streaming applications [Thakur et al. 2012; Usman et al. 2016; Wolf and Pinson 2009]. Similarly, hyperspectral imaging systems stream data from the camera to a computer that controls the system and stores data throughout the acquisition. Dropped frames can occur in hyperspectral imaging systems for several reasons, spanning hardware and software issues.

Dropped frames have been reported in relation to multiple hyperspectral systems [Abd-Elrahman et al. 2011; Chamberland et al. 2004; Guerin et al. 2011] and may not be visible in the final image, especially if only one

image is acquired of a target. Other methods, such as Fourier transform hyperspectral imaging, which records interference patterns over time, work in different ways to line-scanning cameras. In these applications, a single dropped frame results in data corruption affecting the whole image, and so readout electronics have been specifically developed to eliminate dropped frames [Chamberland et al. 2004]. A different approach is taken in an airborne system where each frame is timestamped to identify and position dropped frames [Guerin et al. 2011]. A software-based approach has also been taken to minimise dropped frames using buffering [Abd-Elrahman et al. 2011]. The latter two approaches [Abd-Elrahman et al. 2011; Guerin et al. 2011] did not exclude dropped frames, but attempt to either identify [Guerin et al. 2011] or minimise [Abd-Elrahman et al. 2011] them. A software-based solution to identify dropped frames after acquisition would eliminate the need for custom electronics and enable the correction of images with dropped frames. Moreover, it would allow for re-processing of existing datasets where no provisions for dropped frames were made.

In high-resolution hyperspectral imaging using a line-scan camera, multiple image strips need to be mosaicked into one seamless composite image cube. If any frames are dropped, then all subsequent pixels in one strip will be misaligned across overlapping areas. Aside from hyperspectral systems, line-scan sensors are used in various applications, including panchromatic and multispectral systems such as the Landsat 8 satellite, which are also shown to drop frames [Micijevic et al. 2011]. In the broader context of line-scan imaging the issue of imputing missing data has been studied, but there is limited literature available on the detection of dropped frames in captured images and current methods rely on dropped frames being identified during the image acquisition [Goswami and Sangeeta 2015].

Previous research to identify dropped frames in videos has implemented genetic algorithms for video streams with 1%–5% dropped frames and achieved varying accuracy between 78%–100%, depending on the number of dropped frames and the test video [Thakur et al. 2012, 2013]. The approach looked to match frames from a video with dropped frames to a reference video without dropped frames. When the number of dropped frames increased, the accuracy declined for most test videos. In other fields, such as heritage imaging, there is a need for highly accurate solutions with less concern for computation speed.

A new approach is needed to identify the location of dropped frames that is accurate, reliable, and characterised for push-broom imaging. This article presents two methods for correcting dropped frames in high-resolution push-broom hyperspectral images. The first method uses a reference image from a separate imaging modality to globally register the hyperspectral frames, thus identifying dropped frames. The second method relies only on the hyperspectral images and aligns vertically overlapping hyperspectral image strips. The two approaches are tested on a simulated dataset, and the corrected images are compared to ground truth images using structural similarity index as a metric [Wang et al. 2004]. The most accurate method is found to be the latter, relying only on overlapping strips, and this has been used to correct for dropped frames in real-world data from a hyperspectral imaging case study. In this study, dropped frames were a particular issue, as the equipment had not been fully tested prior to imaging. After the system had been commissioned, dropped frames were less of an issue but could still occur, as reported in other line-scan systems. A robust method for correcting dropped frames remains valuable.

2 METHODS

Correcting dropped frames can be presented as an optimisation problem. Given a frame-wise similarity metric, we are aiming to find the position for each frame in an image that maximises the similarity between either a reference image or overlapping section of two images. The two approaches proposed in this work use the A* algorithm to find the best position for frames while maintaining the sequential order in which they were recorded. We compare the A* algorithm with and without a heuristic function (Section 2.1) and also apply the A* algorithm in two ways: comparing to a reference overview image (Section 2.3) and comparing overlapping strips (Section 2.4). The A* algorithm is a graph algorithm able to find the shortest path between two nodes [Hart et al. 1968]. Here it is used to match frames from two images. In the first approach, frames from the hyperspectral image

(a) Reference image available in Matlab test image library (Cameraman).

(b) Reference image with randomly selected frames to remove.

(c) Reference image with selected frames removed.

Fig. 2. The process of simulating dropped frames.

are matched to a reference image. The second approach matches the frames from two vertically overlapping hyperspectral image strips to each other without using a reference image. Where dropped frames are identified they are filled using 2D interpolation and a binary mask is also produced. More advanced in-fill methods, which better preserve the spectral integrity of the image, have previously been investigated [Goswami and Sangeeta 2015] and are beyond the scope of this article.

2.1 A* Algorithm

The application of the A* algorithm is introduced here on the "cameraman" test image with simulated dropped frames; this is a 256 x 256 pixel grey-scale image that has previously been used as a test image in other applications [Solomon and Breckon 2011; Thyagarajan 2011]. Columns of pixels have been randomly removed from the image to simulate dropped frames. This process is shown in Figure 2.

In this example, normalised mutual information is used as a similarity metric. The normalised mutual information distance between two greyscale frames a, b is defined as $M(a,b) = 1 - I(a,b)/\sqrt{H(a)H(b)}$, where $I(a,b) = H(b) - H(b|a)$ is the mutual information between a and b, $H(f_i) = -\sum_{i=1}^{n} P(f_i)\log_2 P(f_i)$ is the entropy of a frame f, and P is the probability mass function approximated by the normalised histogram of the frame [Strehl and Ghosh 2002].

A cost matrix is calculated using the normalised mutual information distance between each individual frame of the input image and each frame in the reference image. Each of these possible pairings of frames is represented by a node in the graph. The nodes are connected to adjacent frames and restricted to an increasing order. For example, the node pairing frame i in the first image with frame j in the second image is denoted $[i, j]$. Node $[i, j]$ is connected to nodes $[i+1, j]$, $[i, j+1]$, $[i+1, j+1]$, and the edges connecting to these nodes have weights defined by the cost matrix. The A* algorithm is used to find the shortest path (i.e., most similar sequence of frames) from the first frame in each image to the last frame in each image. The two images could be either a co-registered reference image or an overlapping section of two push-broom images.

An example of a cost matrix for a 10-frame image with 3 frames removed is shown in Figure 3. In this example, frames 4, 7, 9 have been removed from the reference image. The optimal path represents the best location for the dropped frames by minimising the total normalised mutual information distance between frames of the two images. Where the optimal path maps one frame in the input image to multiple frames in the reference image, dropped frames are identified. One mapping is the true location of the input frame, and the additional mappings for the input frame are where frames have been dropped from the input image. To determine the true location (and therefore the dropped frames by elimination), the mappings are ranked in decreasing cost and the lowest cost mapping identified as the true location for the input frame.

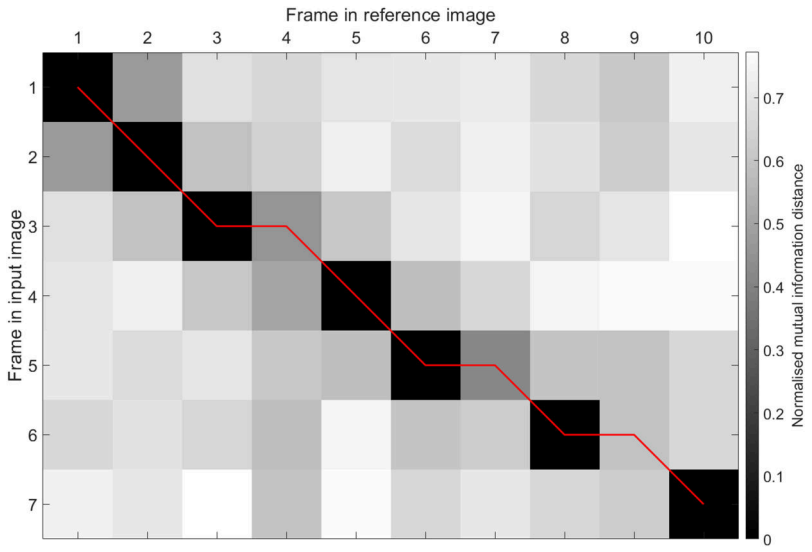


Fig. 3. A distance matrix calculated from the normalised mutual information distance between frames. The optimal path through the distance matrix is shown in red. Darker squares represent a better match with a lower normalised mutual information distance.

2.1.1 Heuristic Function. The objective function of the A* algorithm has two components. The first is the calculated cost to get from the start node to the current node and the second is a heuristic function to estimate the cost from the current node to the goal node. When the heuristic function is not used, the algorithm behaves as Dijkstra's algorithm, which is guaranteed to find the optimal path [Dijkstra et al. 1959]. With large images, however, Dijkstra's algorithm requires significant computational resources, as we will show.

Using a heuristic function can reduce the computational complexity of the algorithm. If a heuristic function is guaranteed to underestimate the shortest distance from the current node to the goal node, then the algorithm is admissible and guaranteed to find the optimal (shortest) path [Hart et al. 1968]. If the heuristic function overestimates the distance from the current node to the goal node, then the algorithm is not admissible and not guaranteed to find the shortest path.

There are many possible heuristic functions and the suitability of each depends on the specific application of the A* algorithm. In many applications the A* algorithm is used with spatial distances [Hart et al. 1968; Tseng et al. 2014]. In spatial applications, a simple and admissible heuristic is Euclidian distance. In this work, the distance metric is not spatial but a measure of similarity between image frames, giving values in the range $[0, 1]$. The shortest possible path is zero if all frames match perfectly, but frames are unlikely to match perfectly in real-world data, which will result in a shortest path greater than zero.

The distance metric used in this work does not use any spatial or ordering information from the frames and depends only on the content of frames across the image. To predict how well a given frame matches across an image would require prediction of the content of other frames across the image. It is therefore not possible to accurately predict how frames will match and instead their similarity (or distance metric) is explicitly calculated.

The empirical heuristic proposed for our work calculates the cost of a theoretical "path" that goes through the minimum values (best matches) of the distance matrix, which are not necessarily connected (i.e., do not necessarily represent a path). Figure 4 highlights the values used in the heuristic calculation for the cost matrix in Figure 3.

If the optimal path goes through all the minimum values, then the heuristic exactly calculates the cost of the path from the current node to the goal node. Otherwise, the path must go through a node with a value greater

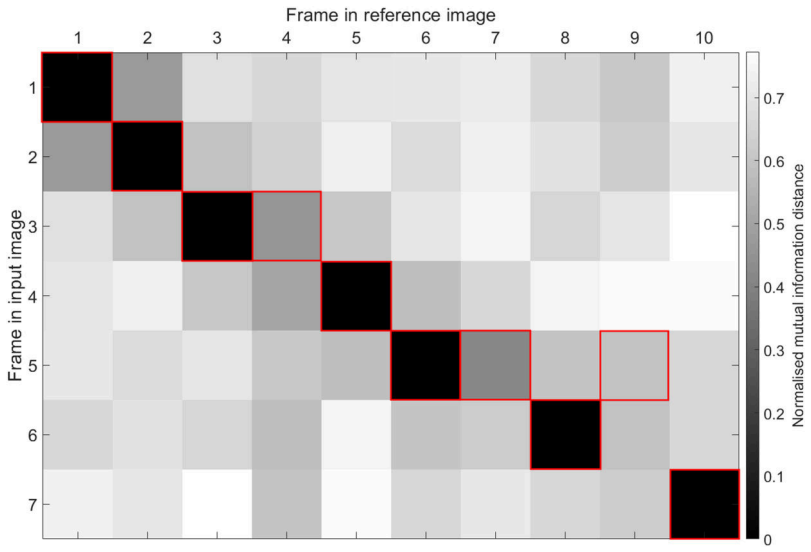


Fig. 4. Showing the values of the distance matrix that are used in the heuristic calculation for an example distance matrix. Darker squares represent a lower normalised mutual information distance and a better match between frames. Note how the resulting "path" would be invalid, as it is not monotonic, but meets the condition of being a lower limit to the optimal cost.

than the minimum, and the heuristic is an underestimate of the path to the goal node. This heuristic is guaranteed to underestimate the shortest path and is therefore admissible.

To assess the computational demand compared to Dijkstra's algorithm, the number of nodes visited by the algorithm is used as a metric. If more nodes are visited, then the algorithm will require more time or computational resources to find the optimal solution. An algorithm that visits fewer nodes is desirable, as it requires less time or computational resources. Dijkstra's algorithm, that is, the A* algorithm with no heuristic, is compared to the A* algorithm with the heuristic proposed above. Both algorithms are run on the cameraman test image (Figure 2(a)) to assess performance against increasing image size and proportion of dropped frames. The results of these comparisons are shown in Figures 5 and 6.

The heuristic function is shown to consistently reduce the computational complexity of the algorithm compared to Dijkstra's algorithm. The heuristic function reduces the number of nodes visited by the algorithm in both cases while finding the same optimal solution.

2.2 Simulated Push-broom Test Images

To simulate a high-resolution push-broom scan the test image is divided into 10 image strips, representing a push-broom scan over a small area of the image. A random number of frames are then removed from random locations in each strip. The cameraman test image is used for introducing these two methods, and both methods are then tested on a set of 12 images in Section 3.2.

Across the 10 strips, the number of dropped frames in each strip is set from a normal distribution with a target mean value x and standard deviation $\sqrt{x}$. Within each individual strip, the frames are removed randomly across the strip with uniform distribution. Five test image sets have been generated from distributions with mean dropped frames $x = 5\%, 10\%, 15\%, 20\%, 25\%$. The test image is divided into 10 overlapping strips, each strip with a height of 35 pixels and an overlapping section of 10 pixels between strips, to simulate one wavelength component of a push-broom hyperspectral image. Figure 7 shows the original reference image and the composite images from strips with varying degrees of dropped frames. The methods presented are used to correct the dropped

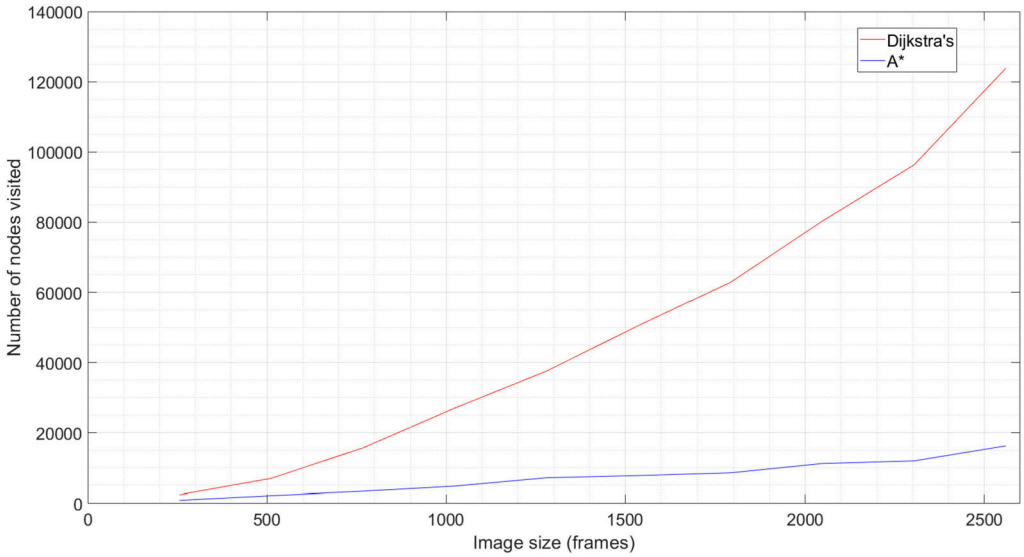


Fig. 5. Graph showing the number of nodes visited as the image size increases. The A* algorithm visits fewer nodes compared to Dijkstra's algorithm.

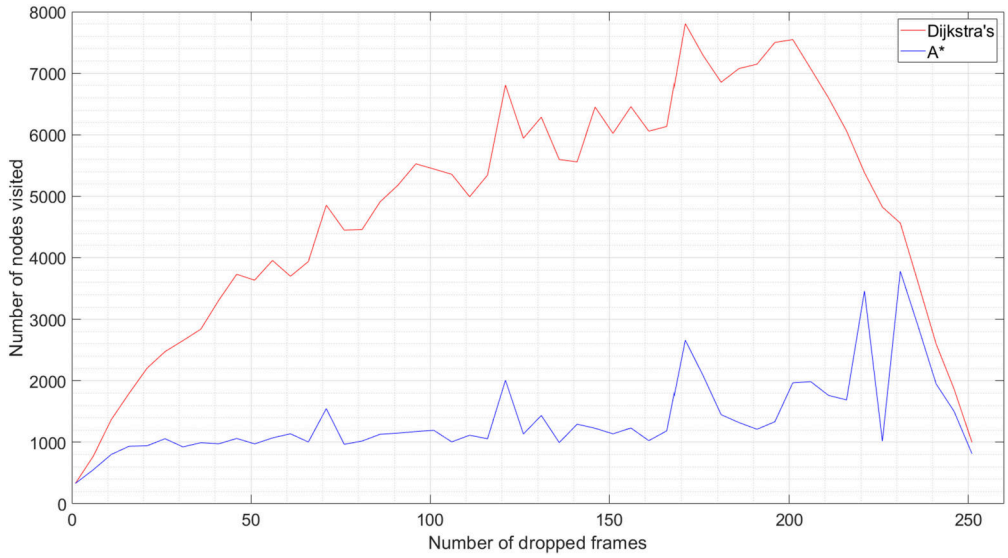
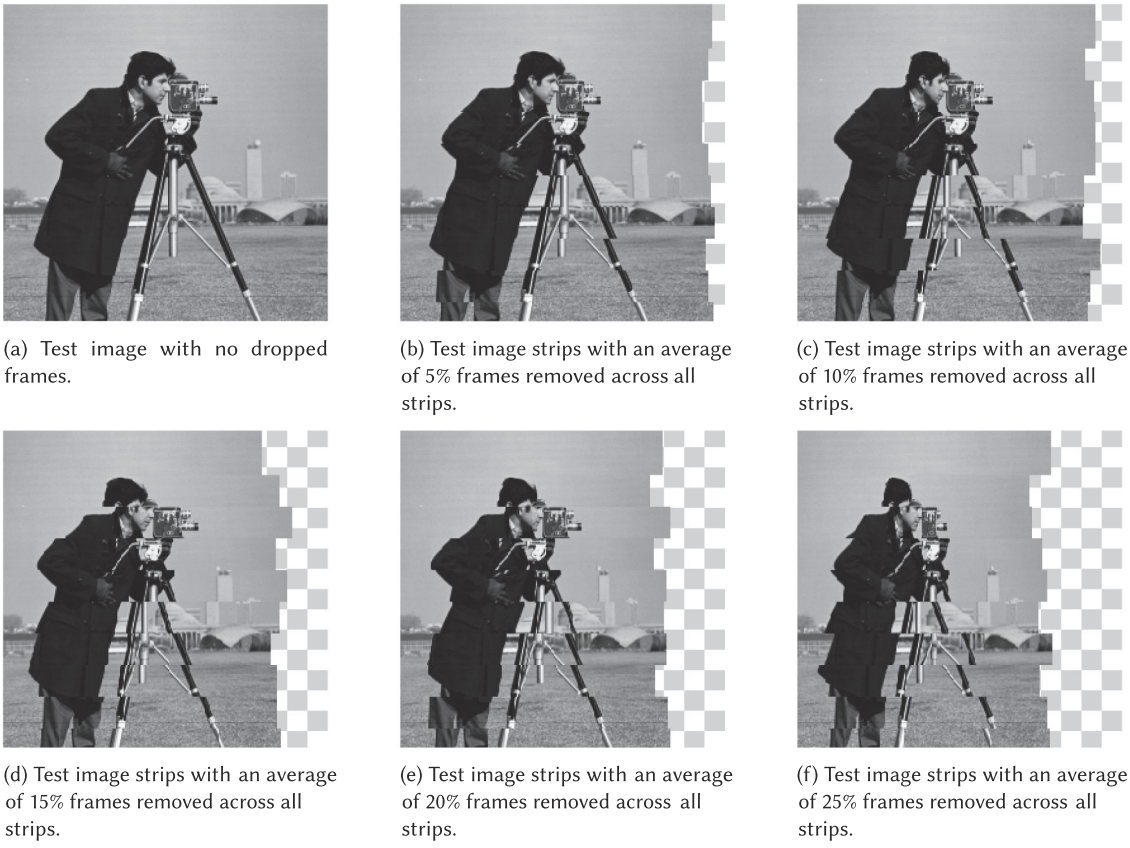


Fig. 6. Graph showing the number of nodes visited as the number of dropped frames increases. The A* algorithm always visits fewer nodes than Dijkstra's algorithm.

frames. To assess the quality of the corrections, alpha blending is applied across the overlapping strips, and structural similarity is calculated between the corrected image and the ground truth image.

2.3 Reference Method

Hyperspectral imaging is often used in conjunction with other imaging modalities. In particular, it is common for an overview image to be available for an imaged scene [Conover et al. 2013, 2015; Rzhanov and Pe'eri 2012].



(a) Test image with no dropped frames.

(b) Test image strips with an average of 5% frames removed across all strips.

(c) Test image strips with an average of 10% frames removed across all strips.

(d) Test image strips with an average of 15% frames removed across all strips.

(e) Test image strips with an average of 20% frames removed across all strips.

(f) Test image strips with an average of 25% frames removed across all strips.

Fig. 7. The test image divided into 10 strips and used to compare both methods with varying proportions of dropped frames.

When overview images are not available from previous studies, images could be taken with a DSLR camera. In high-resolution hyperspectral imaging, this reference image can be used to align multiple hyperspectral image strips and to provide a reference that can be used to identify dropped frames in the hyperspectral image. In general, the reference image and hyperspectral image will have different resolutions, **signal-to-noise ratios (SNR)** and spectral ranges. It is also likely that the overview image will be taken at a different resolution and perspective than the hyperspectral image. This method aims to minimise the normalised mutual information distance between each frame of the hyperspectral image and the overview image. For this calculation, both images need to be converted to greyscale values. An approach that best displays similar features in both the hyperspectral and reference images is desired. In cases where the spectral range of the reference image overlaps with the spectral range of the hyperspectral image, it is possible to compare the two images in the same spectral range. Where this is not possible, a single wavelength band could be selected from the hyperspectral image that achieves the highest mutual information with the overview image. After converting both images to greyscale values, each frame is represented by a single column of pixels. An example of a push-broom image strip with dropped frames is shown in Figure 8, and an example of a reference image strip, cropped from a larger overview image, is shown in Figure 9.

This method requires the reference image to be registered to the hyperspectral images. This is non-trivial, as the hyperspectral images have dropped frames, which will distort the images along one spatial dimension.

Fig. 8. Example input strip with dropped frames.

Fig. 9. Example reference strip.

Fig. 10. Example input strip with dropped frames.

Fig. 11. Example input strip with dropped frames, overlapping the previous strip.

Further to this, the reference image will be from a separate modality that may have taken the image under a separate perspective, with lens distortion and different spectral/radiometric sensitivity. These factors make it difficult to achieve a perfect registration. Even if accurate registration is possible, it will require transformation of the reference image. To assess the robustness of this method, scale and rotation transformations of various magnitudes have been applied to the reference image, and results are discussed in Section 3.

2.4 No-reference Method

In high-resolution hyperspectral imaging, multiple image strips are acquired to achieve high spatial resolution. The individual image strips overlap to allow for alignment and blending between strips. The overlapping sections share the same spatial resolution and the same image plane. As both images are taken with the same modality, Pearson correlation is used as a metric across the overlapping frames. A simulation of two overlapping push-broom image strips with dropped frames is shown in Figures 10 and 11.

In this implementation, the distance matrix is calculated across the overlapping sections of the image strips. The optimal path shown in Figure 12 is used to identify dropped frames in both image strips.

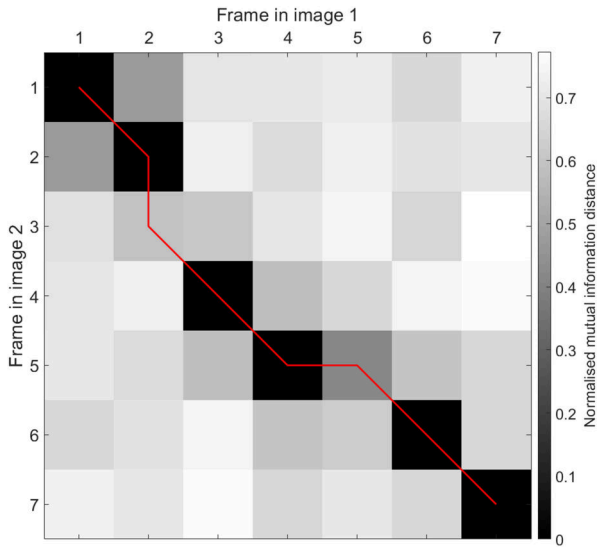


Fig. 12. Showing the optimal path through the distance matrix of two overlapping strips. Each image contains dropped frames independent of other strips.

In this example, the second frame in image 1 matches with the second and third frames in image 2; this suggests that image 1 has dropped its third frame. The fifth frame in image 2 matches with the fourth and fifth frames in image 1. This suggests that image 2 has dropped a frame in this location. It is possible that both images have dropped a frame in the same location and that this will not be detected. In a case with multiple overlapping strips, a frame dropped from both images could be detected in the second overlapping section. When there are multiple overlapping image strips, there is a solution for each overlapping pair of strips that aligns them with one another. With the exception of the first and last images, each strip has two overlaps, one at the top and one at the bottom. To combine the results from each overlap into one image, the positions of frames are propagated through the image strips. The first overlap may introduce dropped frames in the second image strip. These dropped frames are added to the image strip and the path through the bottom overlap. This process is iterated for each overlapping section working from the top strip to the bottom strip, then also in reverse from the bottom strip to the top strip. This ensures that the alignment between each overlapping section is maintained, but may also result in an image that is wider than the ground truth or expected result. To reduce the width of the image to the expected value, the columns of the corrected composite image are ranked in decreasing order of the number of dropped frames they contain. The highest-ranking columns (with the most dropped frames) are removed from the image until the width is the expected value.

3 RESULTS

The reference method has been assessed with various transformations applied to the reference image, which are then reverted. For example, the reference image was downsampled to a lower resolution and then upsampled to match the resolution of the push-broom image to represent the transformation of a low-resolution reference image to align with a higher resolution hyperspectral image. This was also done with rotations applied to the downsampled reference image to simulate a difference in resolution and perspective of the reference image relative to the hyperspectral image while assuming that perfect registration is possible. This was run on the cameraman test image set with an average of 5% dropped frames (Figure 7(b)).

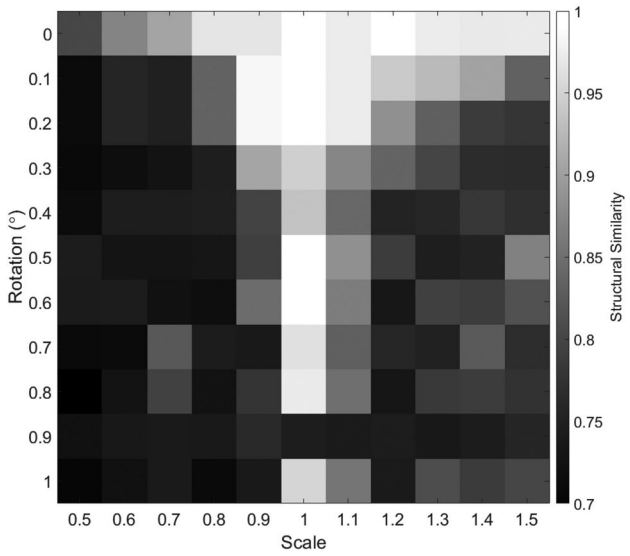


Fig. 13. Showing how differences in rotation and scale between the reference image and image strips can reduce the quality of the correction. Lighter squares represent better results, and darker squares show a worse result.

Figure 13 shows that changes in rotation and scale reduce the quality of the result and the technique is not robust to rotation and scale differences between the reference image and hyperspectral images. If the technique was robust, then the results would maintain a high structural similarity score across changes in scale and rotation and Figure 13 would show little variation.

In the following simulations, two representative scenarios are assessed for the reference method. One where the reference image has no scale or rotation difference to the hyperspectral images, representing a perfect reference image. A second case is also used where the reference image is half the resolution with a one-degree rotation relative to the hyperspectral images, which is chosen to represent real-world applications for high-resolution hyperspectral imaging applications where high-resolution and perfectly aligned reference images are not easily acquired.

3.1 The Impact of the Number of Dropped Frames

The two methods are compared with increasing proportions of dropped frames across the image strips. Figure 14 shows that the performance of the no-reference method deteriorates as the proportion of dropped frames increases. The no-reference method gives a better result than the realistic reference scenario for dropped frames between 5%–20%. When 25% of the frames are dropped, both the no-reference and realistic reference methods give a similar result. The result for the perfect reference method decreases, as there are more frames missing from the input image.

3.2 Other Test Images

A set of 12 test images (Figure 15) have been selected to test the performance of bothmethods across different image features. Images with predominantly vertical and horizontal features were tested as well as images relevant to heritage applications. All images were scaled to a width of 256 pixels. Each image was divided into 10 strips, and an average of 10% of frames were randomly removed from each image strip. The reference method was run

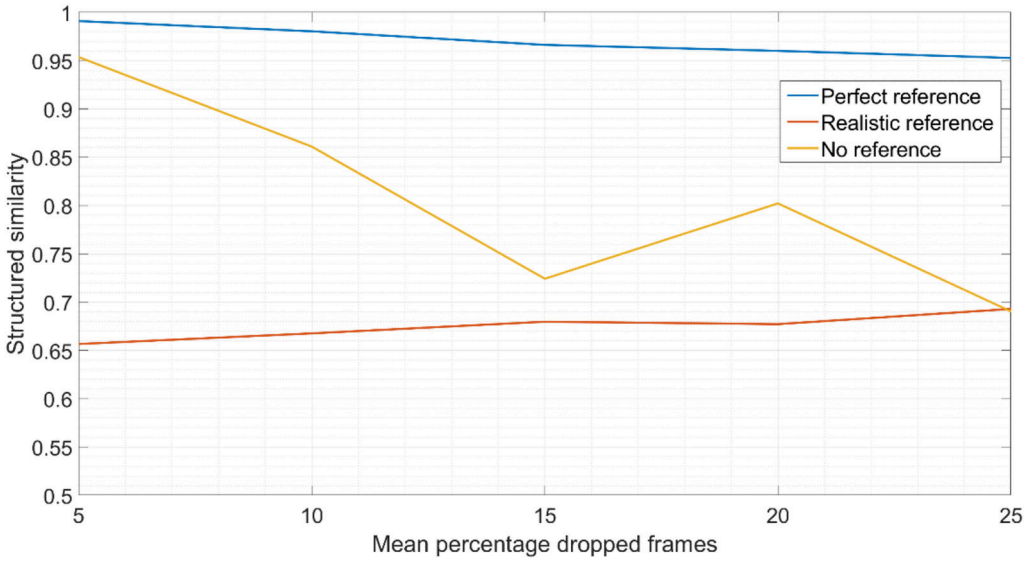


Fig. 14. Showing the quality of the correction of dropped frames for reference and no-reference methods against varying levels of dropped frames.

with a perfect reference and with a half-resolution one-degree rotation applied to the reference; these methods were compared to the no-reference method and the results are shown below.

Figure 16 shows that for all test images the perfect reference method gives the best result. The no-reference method outperforms the realistic reference method on 8 out of the 12 test images. The no-reference method performs worst on the checkerboard, horizontal, and vertical stripes and Composition A test images. This is likely due to the lack of features in the overlapping sections.

4 DISCUSSION

If it is possible to acquire a reference image with low distortion and to perfectly register this to the hyperspectral images, then the reference method achieves the best result. In practice, it is challenging to obtain such a reference image at high spatial resolution, and small differences in resolution or rotation are shown to reduce the quality of the result from the reference method. The no-reference method provides a better result in the realistic scenario with a reference image at half the spatial resolution and with a small one-degree rotation relative to the push-broom images. When processing hyperspectral images, the no-reference method could be implemented with a metric that utilises the spectral information in the overlapping sections of image strips. It is also possible that a larger overlapping section could improve results.

5 CASE STUDY

Dante Gabriel Rossetti painted *La Ghirlandata* in 1873. The oil on canvas painting, measuring 124 × 87 cm, is in the City of London Corporation collection held at the Guildhall Art Gallery, where it has recently undergone full conservation treatment funded by the Bank of America Art Conservation Project [Dex 2018]. Before conservation work began, UCL was given the unique opportunity to image the painting using a newly installed hyperspectral imaging system designed and manufactured by ClydeHSI. Given the time sensitivity of the opportunity, the imaging study proceeded before the hyperspectral system had been fully tested and commissioned. A significant number of dropped frames were found to occur across the hyperspectral images. This has since been

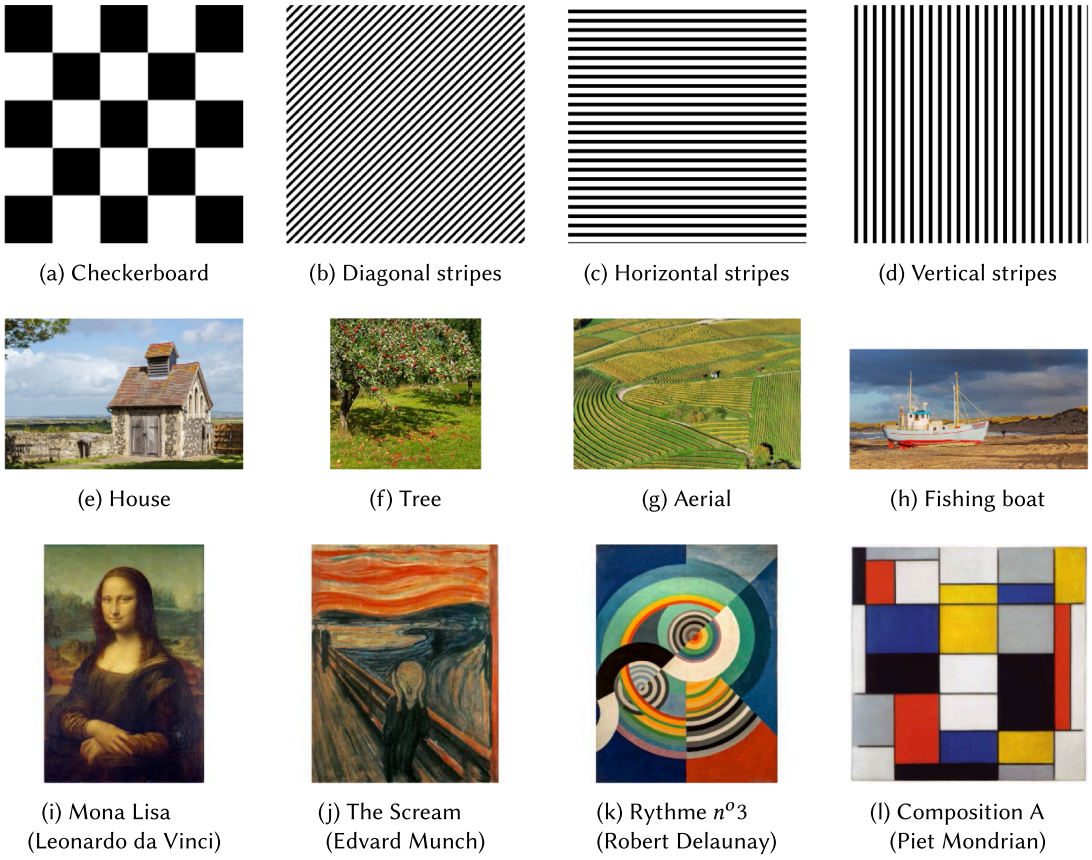


Fig. 15. A set of 12 test images with varying features. Images (e)–(h) are licenced under Creative Commons and GNU licences (CC BY-SA 4.0, CC0 1.0, GFDL-1.2-or-later, CC BY 3.0), respectively. Images i– l are public domain via Wikimedia Commons. [e: © Kim Hansen https://bit.ly/32aUuSI, f: © W. Carter https://bit.ly/2GHqAwZ, g: © Wladyslaw Sojka https://bit.ly/3jVT7x4, h: © Slaunger https://bit.ly/2GIm2GH, i: https://bit.ly/2F1rBzC, j: https://bit.ly/3bFlgph, k: https://bit.ly/3jVSymX, l: https://bit.ly/2F9NCMy].

rectified using hardware and software modifications during the commissioning of the system. The painting was imaged using two hyperspectral cameras; one was a **visible and near-infrared (VNIR)** camera manufactured by ClydeHSI, sensitive to wavelengths 400–1,000 nm at a spectral resolution of 3 nm. The second camera was manufactured by Specim and is sensitive to **short-wave infrared wavelengths** (**SWIR**, 1,000–2,500 nm) at a spectral resolution of 6 nm (Table 1). The VNIR camera acquired 52 image strips across the painting at a spatial resolution of 80 µm/pixel. The SWIR camera acquired 48 image strips across the painting at a spatial resolution of 94 µm/pixel. The cameras were individually mounted on a linear stage and scanned horizontally across the painting, acquiring multiple image strips. At such high spatial resolution, the field of view of the hyperspectral camera is small, and each image strip covers the width of the painting with a small vertical field of view. Each image strip was acquired at a different vertical position with an overlap of 15% (40 pixels) between consecutive strips. The image strips need to be combined into one composite image forming the final composite hyperspectral image.

The image strips exhibit varying numbers of frames recorded across the width of the painting, indicating that dropped frames had occurred. The expected number of frames was 10,875 with the VNIR camera and 9,255 with the SWIR camera, but in fact between 10,810–10,650 and 9,188–8,399 were recorded, as seen in Figures 17 and 18.

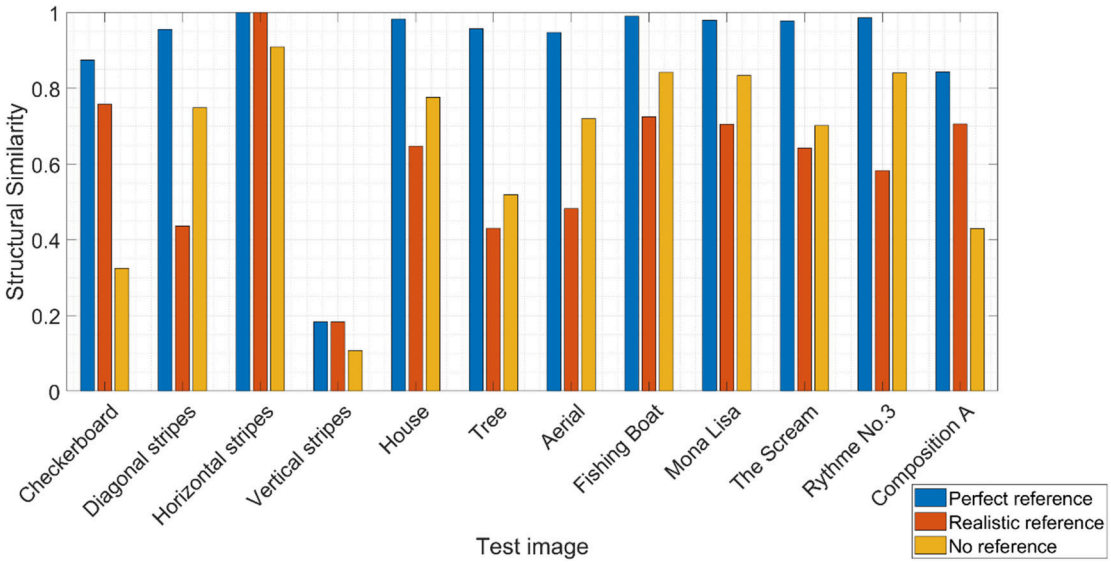


Fig. 16. Showing the quality of results from reference and no-reference methods across a range of 12 test images.

Table 1. Hyperspectral Camera Specifications and Scan Parameters

Parameter	VNIR	SWIR
Spatial resolution [μm]	80	94
Spectral resolution [nm]	3	6
Spectral range [nm]	400–1,000	920–2,500
Lens focal length [mm]	50	56
Aperture [$f/\#$]	5.6	2
Exposure time [ms]	4, 6, 8	2
Working distance [mm]	250	300
Scan speed [mm s^{-1}]	5	2.94
Frame rate [fps]	61.7	31
Vertical pixels [pixels]	324	320
Number of strips recorded [#]	52	48

Overview images from previous imaging studies by Art Analysis & Research were made available via Guildhall Art Gallery. These included a visible-light RGB image and an infrared image taken with an InGaAs sensor. These images had a spatial resolution of 145 μm/pixel and a rotation of approximately 2 degrees relative to the hyperspectral images. Based on the results in Section 3, the best results would be achieved by the no-reference method.

Images from the SWIR camera showed the most variation in the number of frames recorded and were most affected (Figure 18). The A* method was applied across the overlapping strips, using the sum of the pixel-wise cosine distances in each frame as a similarity metric that utilises the spectral information in the overlap. The resulting image strips have been blended using Poisson blending across each wavelength to produce a seamless composite hyperspectral image [Pérez et al. 2003].

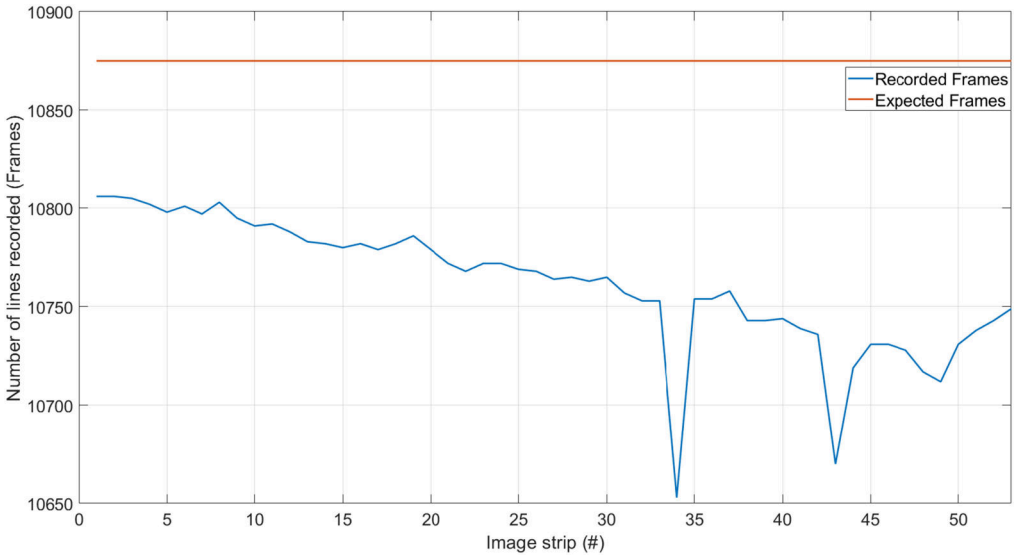


Fig. 17. The number of frames recorded for each VNIR image strip across the painting.

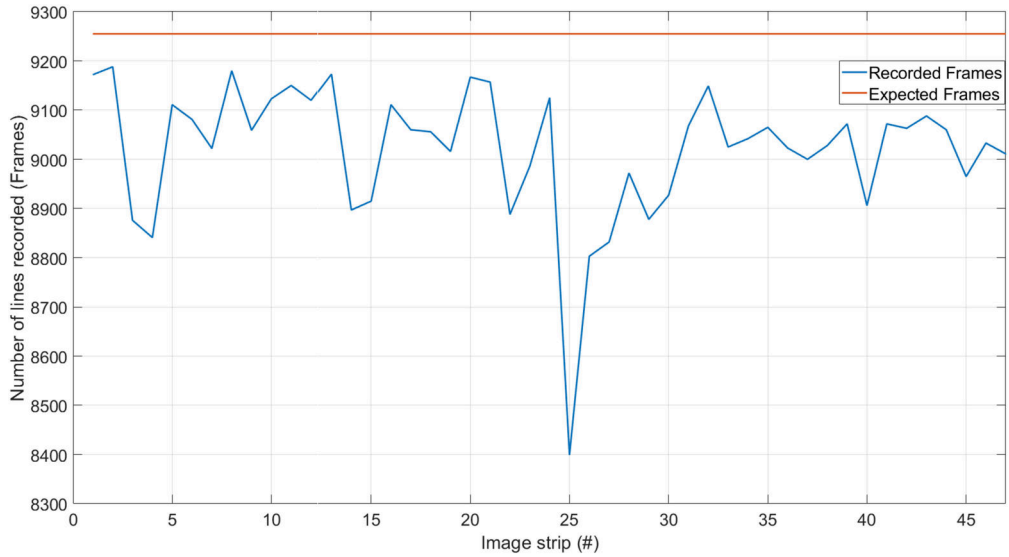


Fig. 18. The number of frames recorded for each SWIR image strip across the painting.

Figures 19, 20, 21 and 22 show uncorrected and corrected SWIR images for visual comparison. The corrected VNIR image is shown in Figure 23. The images have a final resolution of 10,875 × 14,697 pixels and a nominal spatial resolution of 80 µm. The dropped frames have been corrected and alignment between image strips has been achieved at high-resolution. This work was able to reveal the small amount of underdrawing present on the painting in more detail; one area of particular interest is a section of handwriting shown in Figure 22. The painting did not present many opportunities for paint sampling due to its relatively good condition, so visual methods of analysis have been very instructive.

Fig. 19. The SWIR image strips with dropped frames fill left-justified.

Fig. 20. The composite SWIR image after filling dropped frames and blending.

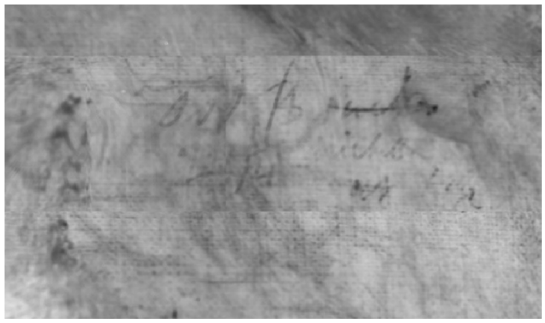


Fig. 21. A detailed section of the SWIR image strips left-justified.

Fig. 22. A detailed section of the SWIR image strips after filling dropped frames and blending.

Analysis of the composite hyperspectral image has suggested pigment identification where paint sampling was not possible and confirmed pigment identification where other methods of analysis had been employed. For example, although all of the figures have similarly coloured blue eyes in visible light, the spectral profile from the eyes of the top right figure is different to the spectra from the eyes of the other two figures in the painting (Figure 24). This suggests that the top right figures' eyes have been painted using a different pigment or mixture. The hyperspectral image also serves as documentation of the painting before conservation and provides the opportunity for further analysis of the painting across visible to short-wave infrared wavelengths.

Fig. 23. The composite VNIR image after filling dropped frames and blending, under simulated D65 illuminant.

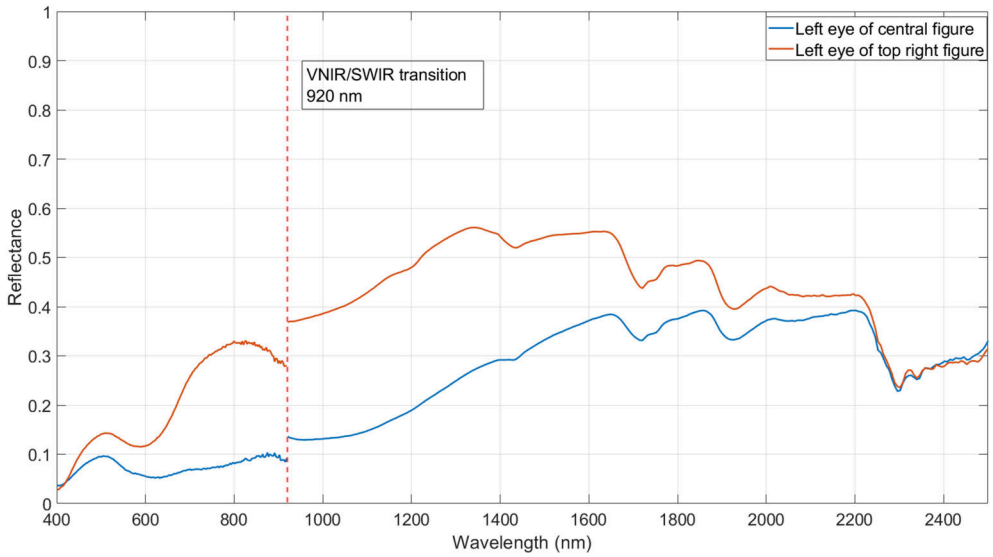


Fig. 24. Spectral profiles of the eyes of central and top left figures.

6 CONCLUSIONS

Two methods have been presented that can correct for dropped frames in push-broom images. One method relies on a co-registered reference image and the other does not require any reference image, but relies on overlap between images to find a globally mutual alignment. The no-reference method performs best in realistic cases where less than 25% of frames are dropped and a reference image of the same resolution and perspective is not available. The reference method performs best if the reference image is the same resolution and can be perfectly registered to the hyperspectral images. If the reference image is well registered but with a difference in resolution and rotation relative to the push-broom images, then the results are often worse than the no-reference method. The no-reference method achieves good results, particularly when less than 10% of frames are dropped. A key advantage of the no-reference method is that it does not require a high-resolution reference image, which makes it more practical in real-world applications of hyperspectral images, particularly across short-wave infrared wavelengths where a high-resolution reference images cannot be easily acquired.

REFERENCES

Amr Abd-Elrahman, Roshan Pande-Chhetri, and Gary Vallad. 2011. Design and development of a multi-purpose low-cost hyperspectral imaging system. *Rem. Sens.* 3, 3 (2011), 570–586.

Andrew Bodkin, A. Sheinis, A. Norton, J. Daly, S. Beaven, and J. Weinheimer. 2009. Snapshot hyperspectral imaging: The hyperpixel array camera. In *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XV*, Vol. 7334. International Society for Optics and Photonics, 73340H.

Martin Chamberland, Vincent Farley, Jean Giroux, Andre Villemaire, Jean-Francois Legault, and Karl R. Schwantes. 2004. Development and testing of a hyperspectral imaging instrument for standoff chemical detection. In *Chemical and Biological Standoff Detection II*, Vol. 5584. International Society for Optics and Photonics, 135–143.

Damon M. Conover, John K. Delaney, and Murray H. Loew. 2013. Accurate accommodation of scan-mirror distortion in the registration of hyperspectral image cubes. In *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIX*, Vol. 8743. International Society for Optics and Photonics, 87431S.

Damon M. Conover, John K. Delaney, and Murray H. Loew. 2015. Automatic registration and mosaicking of technical images of Old Master paintings. *Appl. Phys. A* 119, 4 (2015), 1567–1575.

Robert Dex. 2018. *Bank of America is Paying to Restore Rossetti's "Very Best Picture."* Retrieved from https://www.standard.co.uk/go/london/arts/rossetti-art-painting-bank-retoration-a4019336.html.

Edsger W. Dijkstra et al. 1959. A note on two problems in connexion with graphs. *Numerische Mathematik* 1, 1 (1959), 269–271.

Patricia A. Favero, Jennifer Mass, John K. Delaney, Arthur R. Woll, Alyssa M. Hull, Kathryn A. Dooley, and Adam C. Finnefrock. 2017. Reflectance imaging spectroscopy and synchrotron radiation X-ray fluorescence mapping used in a technical study of The Blue Room by Pablo Picasso. *Herit. Sci.* 5, 1 (2017), 1–18.

Francesca Gabrieli. 2021. Technical aspects and data processing of reflectance imaging spectroscopy for operation night watch. In *Optics for Arts, Architecture, and Archaeology VIII*, Vol. 11784. International Society for Optics and Photonics, 1178414.

Nahum Gat. 2000. Imaging spectroscopy using tunable filters: A review. In *Wavelet Applications VII*, Vol. 4056. International Society for Optics and Photonics, 50–64.

Soumya Goswami and K. Sangeeta. 2015. Anomalies in landsat imagery and imputation. In *Proceedings of the 3rd International Symposium on Women in Computing and Informatics*. 353–358.

Daniel C. Guerin, John Fisher, and Edward R. Graham. 2011. The enhanced MODIS airborne simulator hyperspectral imager. In *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVII*, Vol. 8048. International Society for Optics and Photonics, 80480L.

Peter E. Hart, Nils J. Nilsson, and Bertram Raphael. 1968. A formal basis for the heuristic determination of minimum cost paths. *IEEE Trans. Syst. Sci. Cyber.* 4, 2 (1968), 100–107.

Esad Micijevic, Ron Morfitt, and Michael Choate. 2011. Landsat 8 on-orbit characterization and calibration system. In *Earth Observing Systems XVI*, Vol. 8153. International Society for Optics and Photonics, 81530E.

Patrick Pérez, Michel Gangnet, and Andrew Blake. 2003. Poisson image editing. *ACM Trans. Graph.* 22, 3 (July 2003), 313–318. DOI: https://doi.org/10.1145/882262.882269

Ingmar G. E. Renhorn, David Bergström, Julia Hedborg, Dietmar Letalick, and Sebastian Möller. 2016. High spatial resolution hyperspectral camera based on a linear variable filter. *Optic. Eng.* 55, 11 (2016), 114105.

Yuri Rzhanov and Shachak Pe'eri. 2012. Pushbroom-frame imagery co-registration. *Marine Geodesy* 35, 2 (2012), 141–157.

Chris Solomon and Toby Breckon. 2011. *Fundamentals of Digital Image Processing: A Practical Approach with Examples in Matlab.* John Wiley & Sons.

Alexander Strehl and Joydeep Ghosh. 2002. Cluster ensembles-A knowledge reuse framework for combining multiple partitions. *J. Mach. Learn. Res.* 3, Dec. (2002), 583–617.

Manish K. Thakur, Vikas Saxena, and J. P. Gupta. 2012. A full reference algorithm for dropped frames identification in uncompressed video using genetic algorithm. *Int. J. Dig. Cont. Technol. Applic.* 6, 20 (2012), 562.

Manish K. Thakur, Vikas Saxena, and J. P. Gupta. 2013. Data-parallel full reference algorithm for dropped frame identification in uncompressed video using genetic algorithm. In *Proceedings of the 6th International Conference on Contemporary Computing (IC3)*. IEEE, 467–471.

Kadayam S. Thyagarajan. 2011. *Still Image and Video Compression with MATLAB.* John Wiley & Sons.

Fan Hsun Tseng, Tsung Ta Liang, Cho Hsuan Lee, Li Der Chou, and Han Chieh Chao. 2014. A star search algorithm for civil UAV path planning with 3G communication. In *Proceedings of the 10th International Conference on Intelligent Information Hiding and Multimedia Signal Processing*. IEEE, 942–945.

Muhammad Arslan Usman, Muhammad Rehan Usman, and Soo Young Shin. 2016. A no reference method for detection of dropped video frames in live video streaming. In *Proceedings of the 8th International Conference on Ubiquitous and Future Networks (ICUFN)*. IEEE, 839–844.

Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli. 2004. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* 13, 4 (2004), 600–612.

Stephen Wolf and M. Pinson. 2009. A no reference (NR) and reduced reference (RR) metric for detecting dropped video frames. In *Proceedings of the 4th International Workshop on Video Processing and Quality Metrics for Consumer Electronics, VPQM*, Vol. 5.

Chenghai Yang, James H. Everitt, Michael R. Davis, and Chengye Mao. 2003. A CCD camera-based hyperspectral imaging system for stationary and airborne applications. *Geocarto Int.* 18, 2 (2003), 71–80.

Received December 2020; revised July 2021; accepted August 2021