

# Journal of Biomedical Optics

BiomedicalOptics.SPIEDigitalLibrary.org

## **Multichannel correlation improves the noise tolerance of real-time hyperspectral microimage mosaicking**

Ryan T. Lang  
Julia Tatz  
Eric M. Kercher  
Akilan Palanisami  
Dana H. Brooks  
Bryan Q. Spring

**SPIE.**

Ryan T. Lang, Julia Tatz, Eric M. Kercher, Akilan Palanisami, Dana H. Brooks, Bryan Q. Spring, "Multichannel correlation improves the noise tolerance of real-time hyperspectral microimage mosaicking," *J. Biomed. Opt.* **24**(12), 126002 (2019), doi: 10.1117/1.JBO.24.12.126002.

# Multichannel correlation improves the noise tolerance of real-time hyperspectral microimage mosaicking

Ryan T. Lang,<sup>a,b</sup> Julia Tatz,<sup>a,b</sup> Eric M. Kercher,<sup>a,b</sup> Akilan Palanisami,<sup>c</sup> Dana H. Brooks,<sup>d</sup> and Bryan Q. Spring<sup>a,b,e,\*</sup>

<sup>a</sup>Northeastern University, Translational Biophotonics Cluster, Boston, United States

<sup>b</sup>Northeastern University, Department of Physics, Boston, United States

<sup>c</sup>Massachusetts General Hospital and Harvard Medical School, Wellman Center for Photomedicine, Boston, United States

<sup>d</sup>Northeastern University, Department of Electrical and Computer Engineering, Boston, United States

<sup>e</sup>Northeastern University, Department of Bioengineering, Boston, United States

**Abstract.** Live-subject microscopies, including microendoscopy and other related technologies, offer promise for basic biology research as well as the optical biopsy of disease in the clinic. However, cellular resolution generally comes with the trade-off of a microscopic field-of-view. Microimage mosaicking enables stitching many small scenes together to aid visualization, quantitative interpretation, and mapping of microscale features, for example, to guide surgical intervention. The development of hyperspectral and multispectral systems for biomedical applications provides motivation for adapting mosaicking algorithms to process a number of simultaneous spectral channels. We present an algorithm that mosaics multichannel video by correlating channels of consecutive frames as a basis for efficiently calculating image alignments. We characterize the noise tolerance of the algorithm by using simulated video with known ground-truth alignments to quantify mosaicking accuracy and speed, showing that multiplexed molecular imaging enhances mosaic accuracy by leveraging observations of distinct molecular constituents to inform frame alignment. A simple mathematical model is introduced to characterize the noise suppression provided by a given group of spectral channels, thus predicting the performance of selected subsets of data channels in order to balance mosaicking computation accuracy and speed. The characteristic noise tolerance of a given number of channels is shown to improve through selection of an optimal subset of channels that maximizes this model. We also demonstrate that the multichannel algorithm produces higher quality mosaics than the analogous single-channel methods in an empirical test case. To compensate for the increased data rate of hyperspectral video compared to single-channel systems, we employ parallel processing via GPUs to alleviate computational bottlenecks and to achieve real-time mosaicking even for video-rate multichannel systems anticipated in the future. This implementation paves the way for real-time multichannel mosaicking to accompany next-generation hyperspectral and multispectral video microscopy. © The Authors. Published by SPIE under a Creative Commons Attribution 4.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JBO.24.12.126002](https://doi.org/10.1117/1.JBO.24.12.126002)]

Keywords: mosaicking; multichannel; real-time; hyperspectral; multispectral; imaging.

Paper 190208R received Jun. 21, 2019; accepted for publication Nov. 14, 2019; published online Dec. 11, 2019.

## 1 Introduction

The development of miniaturized devices for live subject microscopy has potential for broad impact on a number of biological research questions and biomedical applications, including the optical biopsy of cancer at the cellular and subcellular levels.<sup>1</sup> Although these devices are typically able to resolve cellular objects, their relatively small field-of-view limits scalability, as clinicians may need to survey large (macroscopic) areas.<sup>2</sup> This limitation stands in the way of practical utility and translation to the clinic. Consequently, achieving both microscopic resolution and macroscopic field-of-view optical imaging remains a central challenge in the field, especially in fluorescence-guided surgery.<sup>2</sup>

Microimage mosaicking offers an approach for expansion of the field-of-view without loss of resolution, aiding visualization and interpretation of microscale features across macroscopic areas of tissue. Mosaicking is an image analysis technique in which sequential frames from a sequence of images (i.e., video) are examined for common spatial features and then stitched

together by overlapping the images' shared regions pairwise.<sup>3</sup> Previous elegant mosaicking algorithms were developed to accompany monochrome (single-channel) microscopy, endomicroscopy, and other related modalities.<sup>4,5</sup> These methods rely only on spatial information to connect frames. These algorithms have since been optimized to address the unique imaging artifacts associated with live-subject microscopy, using creative image alignment and stitching methods.<sup>6,7</sup>

The advent of multidimensional imaging techniques, such as multispectral and hyperspectral imaging,<sup>8-10</sup> and their translation to endoscopy and other multiplexed molecular imaging techniques,<sup>11-13</sup> provides novel information about frame-to-frame motion that single-mode techniques are not equipped to analyze. For example, the existence of strong narrow-band signals from multiple fluorophores can be leveraged to solve the matching problem in a higher signal-to-noise ratio (SNR) domain after unmixing,<sup>14</sup> simultaneously improving accuracy and computational efficiency. Thus, existing mosaicking algorithms should be adapted to leverage spatial and spectral data and enable these techniques to image micro- and macroscopic environments with greater fidelity than their single-channel predecessors.

This work presents a first step in the development of multichannel mosaicking algorithms. We focus on characterizing and

\*Address all correspondence to Bryan Q. Spring, E-mail: [b.spring@northeastern.edu](mailto:b.spring@northeastern.edu)

measuring the improvements to mosaic image registration and alignment using spatial information across two or more channels and comparing them to analogous single-channel methods. Here, we use multichannel in reference to either the set of raw spectral bands produced by hyperspectral imaging or the set of endmember abundance maps calculated via spectral unmixing. We present an application of this work using hyperspectral imaging in a biological context. However, the algorithm may be useful in a broad range of applications where multichannel methods are employed, such as remote geological sensing and astronomical imaging, if used in a setting where the dominant component of motion is translational and any nontranslational motion is small in comparison (see Sec. 3.3).

When processing multiple spectral (spatially coinciding) channels, it is clear that a single unique spatial alignment should be applied across all channels in a set of hyperspectral frames (provided achromatic optical performance). Independent alignment of different channels could obscure colocalization information, which is often of primary interest in hyperspectral data analysis.<sup>10,15</sup> By calculating alignments independently for several spectral channels and leveraging this extra information, we demonstrate that it is possible to improve frame alignment accuracy and tolerance to noise beyond the performance metrics set by their single-channel counterparts.

The inclusion of spectral information into the image alignment calculations adds nuance to the method and is explored in this work. Specifically, there is a trade-off between the addition of spectral channels for increased accuracy and computational speed, from which the user should be able to determine the optimal configuration for their data set. First, some channels within a multichannel data set may be more valuable to the mosaicking accuracy than others depending on their degree of correlation with each other. That is, using totally uncorrelated channels (e.g., two spectrally distinct fluorescence channels) increases the noise tolerance of the mosaic more than the addition of correlated channels. Therefore, a subset of the higher priority channels may be selected to reach a target accuracy and speed. Informed by this, we present a method to choose the optimal subset of channels that maximizes the mosaicking performance. We show that this new metric, called the dimensionality score ( $D$ ), reliably predicts the noise tolerance of both hyperspectral and unmixed multichannel data sets. Therefore, the highest priority channels are predicted to be those that maximize this metric.

Previous reports simulate video registrations from a single image to characterize mosaicking accuracy against a known ground truth.<sup>16</sup> Here, we employ a similar method to characterize the noise tolerance and computational efficiency of the presented algorithm to enable quantification of mosaicking performance relative to single-channel analogues. Simulated hyperspectral video frames were generated with a constant, high percentage (94%) area overlap between consecutive frames to approximate conditions for video-rate imaging. The current algorithm is primarily intended for handheld probes, for which probe motion during acquisition is unavoidable. To combat motion artifacts, these probes are operated at fast imaging rates (video-rate speeds) that outpace the probe motion during image acquisition. The nonoverlap area percentage is simply the ratio of probe displacement (i.e., in pixels traveled) to image size, which must be kept small as a direct consequence of this method of minimizing motion artifacts. Therefore, while a low amount of overlap is common for microscopy tilings or similar methods,

this algorithm is intended for use with systems that produce sequential images with large overlapping areas.

Noise was then introduced to the simulated frames to stress the algorithm and introduce inaccuracies. Additive Gaussian, multiplicative Gaussian, and Poisson noise types were investigated, covering common models that approximate dominant noise sources in both confocal microscopy and hyperspectral imaging.<sup>17,18</sup> We apply the simulated noise tests to quantify the improvements made by the present multichannel mosaicking algorithm as well as to measure the predictive power of the aforementioned dimensionality score. These hyperspectral images were also linearly unmixed into their molecular constituents, creating unmixed channels for which  $D$  was calculated. However, note that unmixed channels are generally denoised compared to spectral channels,<sup>14</sup> which is intrinsically superior to using raw spectral channels but relies on faster than real-time unmixing for online applications.

Developing a precise quantification of mosaicking error to measure the noise tolerance requires objective quantification that may be beyond human perceptions of mosaicking errors and artifacts. To ensure that the improvements found for the noise tolerance translate into improvements in image quality as perceived by the eye, we developed a separate empirical test. We collected a second set of hyperspectral frames using a commercial microscope and compared mosaics made from our algorithm to gold standard tiling of the same area assembled using the microscope's precision stage and software.

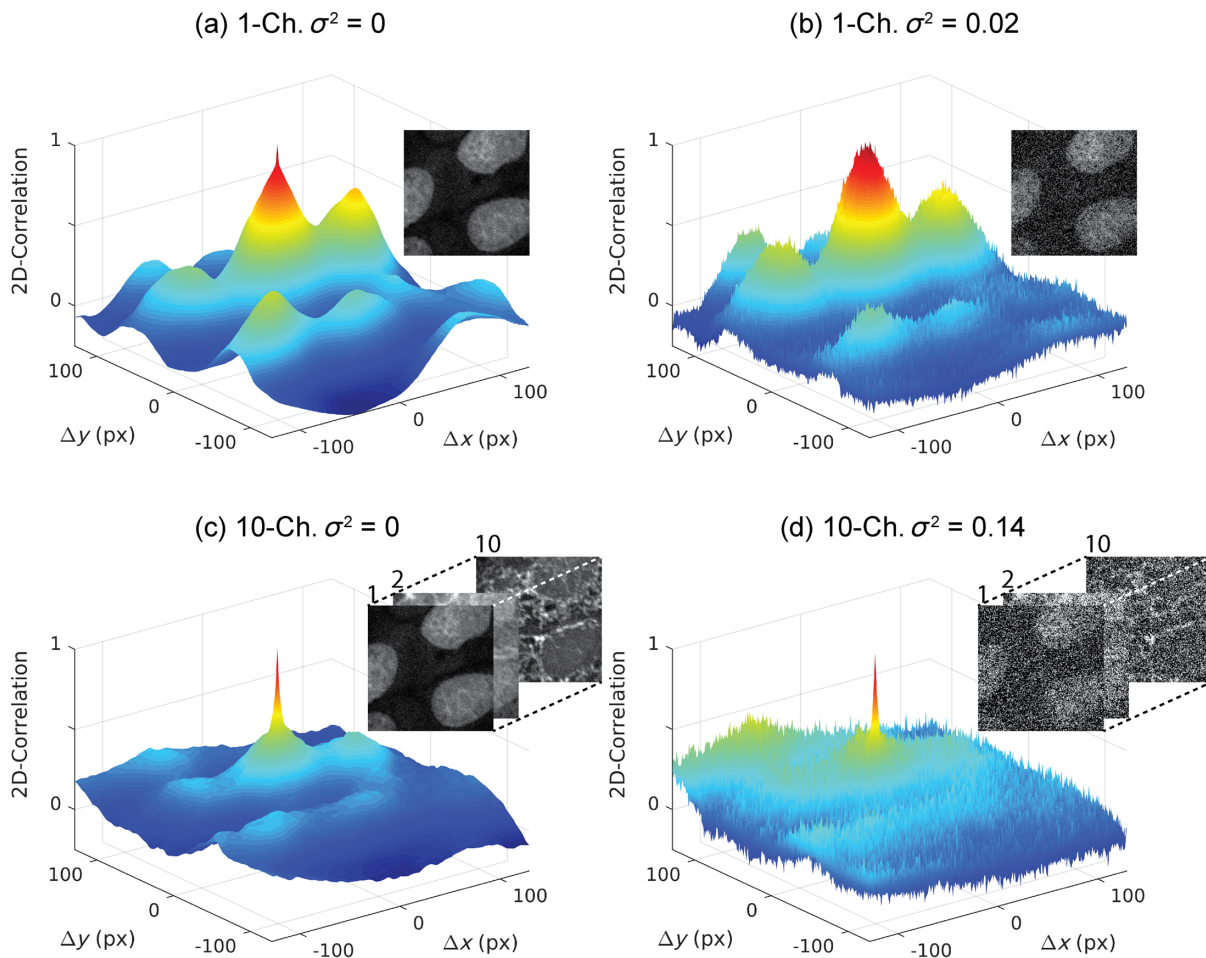
Furthermore, we leverage improvements in computational throughput available with parallelized computing and graphics processing units (GPUs),<sup>19</sup> allowing computationally expensive multichannel analysis at video-rate (15 frames-per-second, fps) speeds for up to 10-channel mosaicking. Thus, image computation speeds can be comparable to associated data acquisition rates,<sup>20</sup> enabling future online applications in the clinic. These improvements may be applied to mosaicking, and pairing these algorithms with the development of faster-than-real-time spectral decomposition<sup>20</sup> would pave the way for real-time construction of macroscopic maps of multiple molecular constituents (endmembers) of tissue with microscopic resolution.

## 2 Methods

### 2.1 Multichannel Mosaicking Algorithm

The presented algorithm applies pairwise normalized cross-correlation<sup>21</sup> (NCC) as a basis for frame registration. Following Bedard et al.,<sup>5</sup> cross-correlation maps are calculated between consecutive image frames [Figs. 1(a) and 1(b)], and the maps are normalized to an in-place autocorrelation (i.e., the theoretical maximum correlation). Correlation space then maps the probability of alignment across the spatial domain spanned by the intersection of two images being aligned, from which the location of maximum likelihood is selected as the "true" alignment location for the subsequent stitching step. We calculate the spatial cross-correlation map separately for each additional channel and average the maps in correlation space, producing a single map [Figs. 1(c) and 1(d)].

The code enabling this algorithm was developed in MATLAB R2019a (The MathWorks) in a Linux environment (Ubuntu 18.04LTS) with an Intel i7 6800K 3.40 GHz CPU and 32 GB DDR4 2400 MT/s RAM. The substantial computational cost of calculating multiple cross-correlations per frame-pair was alleviated by using MATLAB's Parallel Computing



**Fig. 1** Exemplary cross-correlation maps demonstrate improved noise tolerance for multichannel versus single-channel micromosaicking. (a) Single-channel NCC map shows typical landscape with global maximum representing location of the correct image alignment. (b) Gaussian noise ( $\sigma^2 = 0.02$ ) added to the images correlated in (a) results in fluctuations in the NCC map that obscure the location of the correct alignment. (c) NCC landscape averaged over 10 spectral channels reduces the probability of an incorrect alignment. (d) The location of the true alignment is not obscured after Gaussian noise [ $\sigma^2 = 0.14$ , a 7-fold increase from (b)] is added to the 10-channel NCC configuration in (c), demonstrating robust noise tolerance compared to the single-channel analog. Image insets provide a visual juxtaposition of noise strength.

Toolbox v7.0 for GPU interfacing. Speed benchmarks were carried out using a GeForce GTX 1080 Ti graphics card (NVIDIA). The code implementing this algorithm is available at the GitHub repository <https://github.com/springlabnu/Multi-Channel-Mosaicking>.

## 2.2 Simulated Data Generation and Mosaicking Performance

Quantitative analysis and comparison to equivalent single-channel methods was carried out using a simulated video-rate data set derived from images obtained by empirical microscopy. We acquired a hyperspectral image cube (26 spectral bands, images shown in Fig S1 in the [Supplemental Materials](#)) of fixed Madin–Darby canine kidney cells stained with three dyes for visualization of the nucleus (Hoechst 33342), cytoskeleton (Alexa Fluor 488-Phalloidin), and mitochondria (MitoTracker Red CMXRos) using a laser scanning confocal microscope (Olympus FV3000). An artificial hyperspectral video stream was then generated from this data by cropping the field-of-view

and traversing the image in a continuous raster scan with 94% area overlap between pairs of images [raster scan pattern illustrated in Fig. S2(a) in the [Supplemental Materials](#)]. This created a sequence of 65 hyperspectral image cubes ( $512 \times 512$  pixels) with a known ground truth alignment for quantitative algorithm characterization. The image cubes were also linearly unmixed into their three molecular constituents, creating three unmixed channels that could also be used for comparative analysis.

Having the ground truth location for each frame means that, for each trial, the alignments may be checked against the lookup table of correct locations without the need to visualize the full mosaic. From here, we quantified mosaicking accuracy  $\eta$  as the fraction of the 64 frame pairs (one less than the number of frames) that were correctly aligned compared to the ground truth (i.e., the original image). The alignment was counted as correct only for an exact match. Zero-mean additive Gaussian noise was added to each image in every frame, in order to degrade image quality and to mimic challenging experiments that could potentially cause alignment mistakes. In order to quantify noise power of the simulated images, we use the peak SNR, defined as

$$\text{PSNR} = 10 \log_{10} \frac{M^2}{\sigma^2}, \quad (1)$$

where  $\sigma^2$  is the added noise variance and  $M$  is the maximal detector readout value of pixels in the image. For this data, the pixel brightness is normalized ( $M = 1$ ) for ease of comparison to  $\sigma^2$ . For example,  $\sigma^2 = 0.2$  corresponds to a variance that is 20% of the peak signal.

The mosaicking accuracy  $\eta$  is approximately equal to the probability that a single pair of frames are correctly aligned [ $P(\text{correct})$ ]; that is, correct alignment is a Bernoulli variable with probability  $\eta$  for which the amount of noise is a predictor. The probability of Bernoulli variables is given by a logistic function;<sup>22</sup> this allows us to connect the fraction of correct alignments ( $\eta$ ) to the noise strength causing the misalignments (PSNR) by

$$\eta \simeq P(\text{correct}) = [1 + e^{-k(\text{PSNR} - \text{PSNR}_{50})}]^{-1}, \quad (2)$$

where  $k$  and  $\text{PSNR}_{50}$  are the free parameters of the logistic fit. While  $k$ , the logistic growth rate, is approximately the same for these trials, the midpoint  $\text{PSNR}_{50}$  (PSNR at which 50% of frames are correctly aligned or where  $\eta = 0.5$ ) is used as a metric of noise tolerance. Finally,  $\sigma_{50}^2$ , the variance corresponding to  $\text{PSNR}_{50}$  [Eq. (1)], is used as a measure of the tolerated noise.

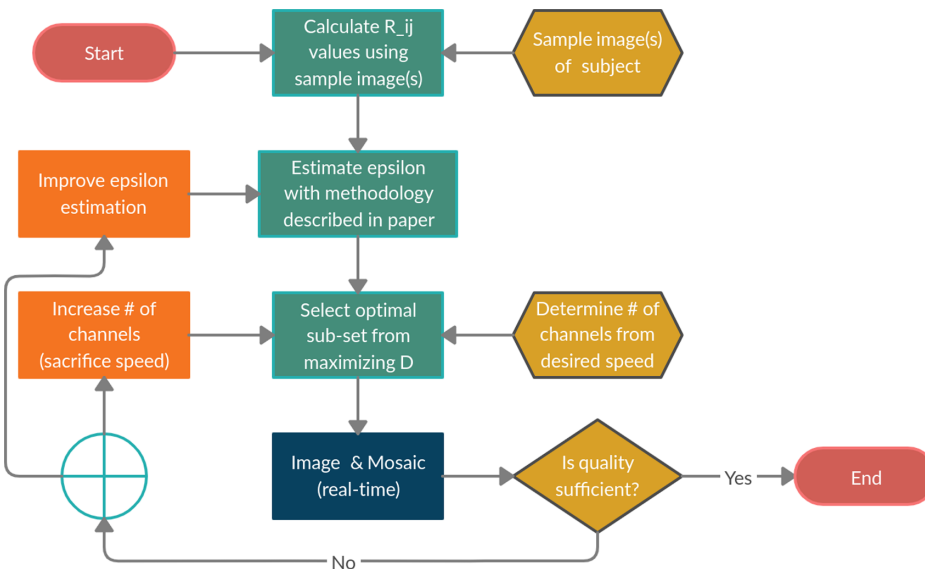
Testing of the present algorithm was carried out using  $n = [2, 26]$  raw spectral channels. The unmixed channels were only analyzed by calculating their  $D$  score to compare their predicted performance. For each  $n$  value,  $D$  was calculated for every permutation of  $n$  channels, from which the permutation with the maximal  $D$  value was selected as the optimal subset. As a control, a permutation was selected subjectively by an operator based on prioritizing channels with the brightest signal corresponding to each molecular constituent, intended to mimic

selection of channels without knowledge of  $D$ . Histograms of all possible  $D$  values (all possible permutations) are shown for these cases in Fig. S3 in the [Supplemental Materials](#) to demonstrate the optimization versus the user-selected permutations. Additive Gaussian noise with  $\sigma^2 = [0.003, 1.000]$  was added to each channel in each image within the configuration, after which the algorithm was run five times, with  $\eta$  being saved after each trial. The time for each trial to complete the full mosaic was also saved. The applied variance was translated to PSNR using Eq. (1), fit to Eq. (2), from which the  $\sigma_{50}^2$  was extracted. Independent two-sample  $t$ -tests (Welch's  $t$ -test) were calculated to find the significance of the improvements in  $\sigma_{50}^2$  through optimization of subset permutation, as well as other comparisons.

These experiments were repeated for multiplicative Gaussian and Poisson noise. For these noise types, the amount of noise power delivered depends on signal, so the  $\sigma_{50}^2$  value extracted is signal dependent as well. The characteristic performance is thus amended to  $\sigma_{50}^2/I^2$  (i.e., normalized by the intensity) accordingly for these noise types.

### 2.3 Empirical Data Generation

While the simulated data analysis thoroughly examines and characterizes the present methods, the ultimate test of the multi-channel algorithm and the predictive power of  $D$  is to verify that they produce useful images from an empirical data set subject to real-data constraints (Fig. 2). To create an empirical data set, ovarian cancer cells (POWDER cell line, Cellaria Biosciences) were plated and stained with two fluorescent dye stains (Hoechst 33342 and DiI) and two fluorescent dye-antibody conjugate markers (PerCP-CD44 and AF647-EGFR). Using a confocal fluorescent microscope with a 32-channel hyperspectral detector (Zeiss LSM880, channel bands shown in Table S1 in the [Supplemental Materials](#)) 30 frames were generated by taking



**Fig. 2** For real data applications, ground truth knowledge is unavailable, requiring users to apply the principles shown here to improve mosaicking accuracy under noise loads without quantitative feedback. Dimensionality scores ( $D$ ) must be calculated from sample frames, which is shown to have some effect on subset optimization (see Fig. S4 in the [Supplemental Materials](#)). In addition, the cutoff value  $\epsilon$  must be estimated using a histogram rather than determined quantitatively. If the channels calculated to maximize  $D$  are not sufficient to prevent unacceptable mosaicking errors, users may include additional channels or attempt to fine-tune the estimation of  $\epsilon$  until results are satisfactory. Note that the  $\oplus$  symbol indicates a logical “or.”

separate images after manual movements of the stage, generating a set of frames with a wide range of overlap amounts (estimated at 77% to 96% based on alignments reported by the algorithm). Mosaicking was carried out for a 10-channel configuration, whose channel constituents were chosen as the set that maximizes  $D$  (i.e., the optimal 10-channel subset). This was compared to a single-channel control representing the analogous single-channel method. The same area was imaged using a tiling protocol available in the microscope's software (ZEN Blue 2.6) to create a gold standard image for qualitative comparison of the mosaics.

### 3 Results and Discussion

#### 3.1 Dimensionality Score

In an effort to predict the optimal subset of channels for multichannel mosaicking, we developed a new metric termed "dimensionality score" that quantifies the degree of independence of a given grouping of data channels, and by extension, predicts the characteristic noise tolerance. The optimal channel set, then, is predicted to be that which maximizes this score (thereby maximizing noise tolerance). This metric must capture the primary means by which the present multichannel algorithm suppresses the effects of noise to retain mosaicking accuracy.

The multichannel integrated NCC map (the mean NCC map; Fig. 1) suggests that the present approach offers two means of noise suppression. Adding noise to the image channels generates similar noise levels but with uncorrelated patterns in the individual correlation maps, and comparison of Figs. 1(b) and 1(d) shows that the variance in NCC space is reduced by taking the mean of multiple NCC maps, since random noise fluctuations are averaged out and reduced over several spectral channels. However, comparison of Figs. 1(a) and 1(c) reveals that this process also populates (sharpen) the central true peak feature while reducing the magnitudes of false peaks. Hence, multichannel mosaicking is more immune to noise compared to the single-channel case not only through averaging out random noise in NCC space but also because this spectral averaging apparently suppresses the false NCC peaks due to separate biological objects that share similar structural features, and therefore enhances the strength of the true peak versus false peaks. The strength of this suppression of false peaks is dependent on averaging over diverse NCC landscapes. Averaging multiple NCC maps from weakly correlated or uncorrelated channels causes true peak enhancement by utilizing a set of independent NCC maps with the same true peak location while stochastic false peaks generally are not amplified and tend to average to zero.

By this reasoning, the noise tolerance of a given set of channels depends primarily on the number of channels and the degree of dependence between those channels. The coefficient of correlation is already a well-known descriptor of the degree of dependence between two samples,<sup>23</sup> so it forms a natural basis for constructing such a descriptor. Since the samples here are images, we can calculate the two-dimensional (2-D) coefficient of correlation  $R_{2D}$ , which is defined for 2-D matrices (images)  $A$  and  $B$  as

$$R_{2D} = \frac{\sum_{m,n} (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{[\sum_{m,n} (A_{mn} - \bar{A})^2][\sum_{m,n} (B_{mn} - \bar{B})^2]}}, \quad (3)$$

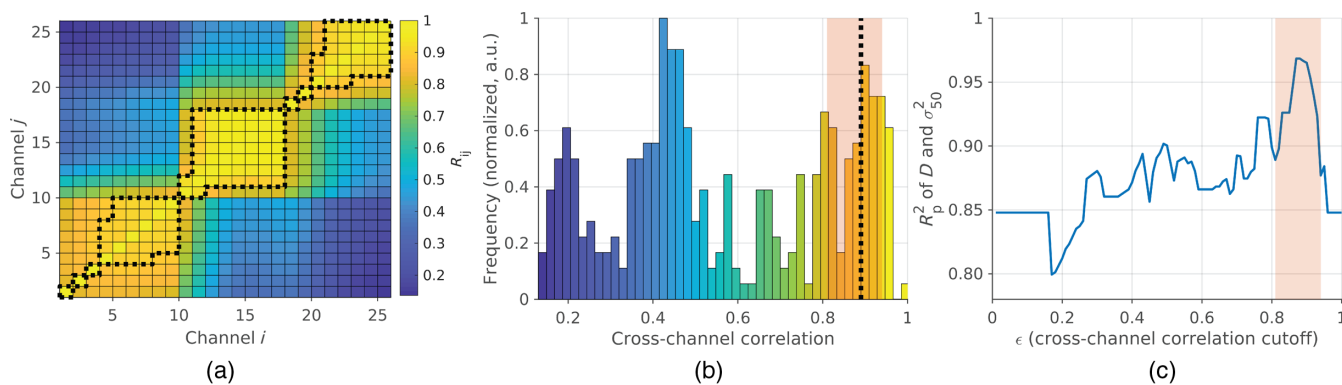
where  $m$  and  $n$  denote rows and columns of the matrix, and  $\bar{A}$  and  $\bar{B}$  denote averages of the matrix elements. This 2-D correlation is computed between pairs of channels  $i$  and  $j$  using the associated ground-truth image for each channel as matrices  $A$  and  $B$  in the formulation above (i.e., matrices  $A$  and  $B$  are given by images from channels  $i$  and  $j$ ). These  $R_{ij}$  values are then separated into two populations based on a cutoff  $\epsilon$ ; strongly correlated channel pairs ( $R_{ij} > \epsilon$ ) have correlation values denoted  $R_C$  and weak correlations ( $R_{ij} < \epsilon$ ) are similarly grouped as  $R_I$ . Hence,  $\epsilon$  acts as a cutoff for deeming channel pairs as either independent ( $R_I$ ) or correlated ( $R_C$ ). The dimensionality score  $D$  is then calculated for  $n$  channels as

$$D = 1 + (1 - \overline{R_I}) * n_{\text{ind}} + (1 - \overline{R_C}) * n, \quad (4)$$

where  $n_{\text{ind}}$  is the number of independent channels (i.e., channels with no correlation values greater than  $\epsilon$ ) and the overbar signifies the mean across each separate collection of channels. The dimensionality score is bounded on the range  $[1, n + 1]$ ; a score of 1 indicates that all channels are exact copies of each other (or at least, are proportional) whereas a score of  $n + 1$  indicates that all channels are completely independent (all  $R_{ij} = 0$ ). The splitting of  $R_{ij}$  into  $R_C$  and  $R_I$  attempts to capture the two means of noise suppression discussed above:  $(1 - \overline{R_I}) * n_{\text{ind}}$  models the effect of true peak sharpening by largely uncorrelated channels and  $(1 - \overline{R_C}) * n$  estimates the contribution of all channels that reduce the noise variance in correlation space through averaging to suppress stochastic fluctuations.

The selection of  $\epsilon$  has a significant effect on the performance of  $D$  as a predictor for  $\sigma_{50}^2$ . An appropriate value of  $\epsilon$  realizes the separation of  $R_{ij}$  values characteristic of channels that image nearly identical features. For a matrix of  $R_{ij}$  values, this effect results in a characteristic division of the  $R_{ij}$  matrix into distinct neighborhoods [Fig. 3(a)]. We demonstrate this effect for  $\epsilon = 0.89$ , leading to three such areas that likely correspond to the three molecular constituents imaged in this dataset. Without the knowledge of a ground truth for grading performance, an appropriate value of  $\epsilon$  may be estimated (Fig. 2) through analysis of a histogram of  $R_{ij}$  [Fig. 3(b)] with the intention of separating the population of strongly correlated channel pairings from the rest. Here, with knowledge of the ground truth, the  $\epsilon$  value that provides the greatest predictive strength may be found by maximizing the Pearson correlation between  $D$  and  $\sigma_{50}^2$  (the characteristic noise tolerance), with a peak value at  $\epsilon = 0.89$  giving a squared Pearson correlation of  $R_p^2 = 0.97$  [Fig. 3(c)]. However, our results indicate that a considerable range [shaded area in Figs. 3(b) and 3(c)] of  $\epsilon$  values correspond to a central peak of Fig. 3(c), and therefore retain near-optimal predictive performance. Therefore, we expect that an estimation of  $\epsilon$  using the aforementioned histogram method will still ensure that  $D$  remains a reliable and useful predictor of mosaicking performance. Note that the limits of  $\epsilon \rightarrow 0$  and  $\epsilon \rightarrow 1$  both reduce  $D$  to  $1 + (1 - \overline{R}) * n$ , which is an intuitive descriptor of dimensionality; simply the mean of one minus cross-channel correlation, scaled by  $n$ . However, in this limit, there is only one operative term ( $\overline{R}$ ). Therefore, it is apparent that both noise suppression modes are not captured in this limit, resulting in a weakened (lower Pearson's correlation) prediction of noise tolerance.

Even with our parallelized implementation, calculating the cross-correlation remains the dominant computational bottleneck, and the mosaicking computation time thus linearly scales with the number of channels used.  $D$  is employed to optimize



**Fig. 3** Selection of the cross-channel correlation cutoff,  $\epsilon$ , affects the predictive accuracy of  $D$  by changing the strictness of “independence” definition. (a) A heatmap of the  $R_{ij}$  matrix shows  $\epsilon = 0.89$  (dotted line) segregating strongly correlated channel pairings into three major areas corresponding to the three molecular constituents imaged. (b) A histogram of  $R_{ij}$  values shows that  $\epsilon$  values within the shaded area separate a population of strongly correlated channel pairings ( $R_{ij}$  values) from the rest (dotted line indicates  $\epsilon = 0.89$ ). (c) For data with known ground truth, predictive power may be optimized directly by maximizing Pearson’s correlation ( $R_p^2$ ) of  $D$  and  $\sigma_{50}^2$  versus dimension score cutoff ( $\epsilon$ ). This reveals an optimal  $\epsilon = 0.89$  corresponding to  $R_p^2 = 0.97$ ; the shaded area from (b) now outlines a central peak, indicating that subprime  $\epsilon$  values estimated from (b) will still deliver near-optimal predictive strength.

the selection of a high-priority subset of channels in an effort to lessen this computational load. The optimal configuration is found by maximizing  $D$  over channel selection choices that are sufficiently speedy (i.e., limiting the number of channels to reach a target mosaicking rate). This enables  $D$  to be a practical tool for optimizing the selection of subsets of available channels. A brute-force calculation of  $D$  scores for all possible permutations of the desired configuration (of which there may be hundreds or thousands) can be computed in a reasonable time frame, providing comparative values that predict each permutation’s noise tolerance. The problem of selecting the optimally performing subset of channels for mosaicking is, to the best of our knowledge, otherwise intractable, except by empirical trial-and-error by the operator combined with a heuristic quality criterion.

### 3.2 Quantitative Mosaicking Performance

Increasing the number of spectral channels improves the mosaicking accuracy under noise load, as indicated by plotting  $\eta$  versus PSNR [Fig. 4(a); see also Eq. (2)]. For a given PSNR, more channels always perform better (higher  $\eta$ ), and at a given  $\eta$ , having more channels allows us to tolerate stronger noise levels (see Table S3 in the Supplemental Materials). These results are replicated for multiplicative Gaussian noise and Poisson noise, as shown in Fig. S5 in the Supplemental Materials. Therefore, the multichannel algorithm is expected to achieve greater accuracy than equivalent single-channel methods for noise models relevant in microscopy. Two permutations were tested for each configuration (number of channels): the optimal permutation predicted by  $D$  and an operator-selected control (these permutations are shown in Table S2 in the Supplemental Materials). The control was assembled by selecting subsets of channels that evenly sampled the three molecular constituents, intended to represent an operator’s educated guess at the optimal configuration of channels (the alternative to using the dimensionality score for optimization). In all cases, the optimized permutation was superior (lower  $\text{PSNR}_{50}$ ) to the control [Fig. 4(b)].

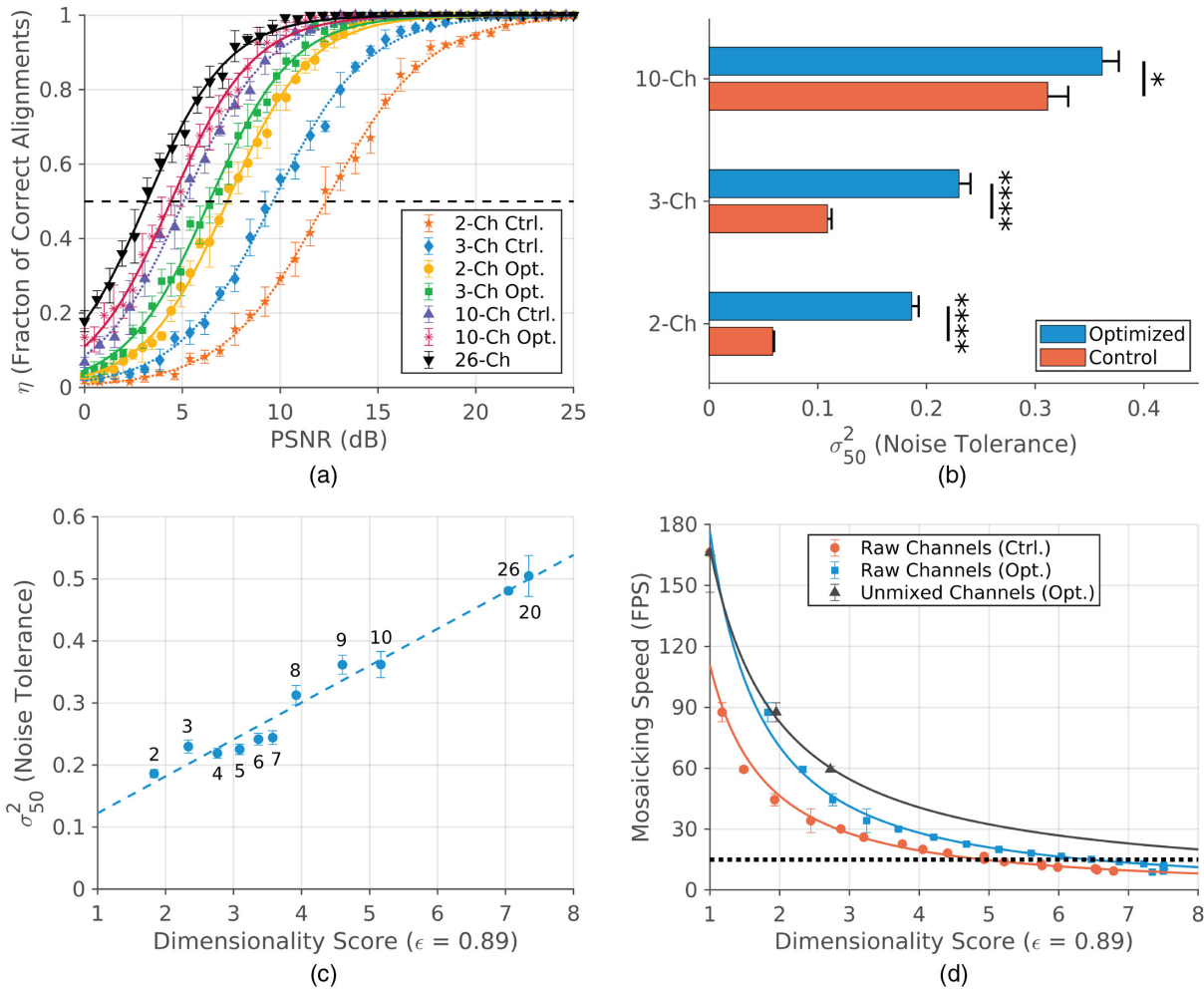
The correlation and thus predictive strength remains high between the  $\sigma_{50}^2$  values and  $D$  [Fig. 4(c),  $R_p^2 = 0.95$ ] for the

optimized subsets. This is very similar to the previous correlation value of 0.97 between  $\sigma_{50}^2$  values and  $D$  for the operator control channel sets. Therefore,  $D$  retains its prediction strength among the various subsets of optimized permutations. In all cases tested, the optimized permutations provide superior performance over the operator-selected ones [Fig. 4(d); at least  $p < 0.05$ ]. While the improvement is most striking for small numbers of channels, the effect is present for all cases, confirming an increase in noise tolerance by use of the dimensionality score to predict the optimally performing subset of channels.

The dimensionality score  $D$  plotted against the mosaicking speed [Fig. 4(d)] demonstrates the trade-off between computational speed (increasing  $y$ ) and dimensionality score (increasing  $x$ ) for both control and optimized permutations of the raw channels, as well as the optimized permutations of unmixed channels. These plots were fit to power law curves, and two-tailed  $t$ -tests were evaluated for the coefficients of these fits to show that: (1) unmixed channels outperform optimized raw channels ( $p < 0.001$ ) and (2) optimized raw channels outperform operator-selected channels ( $p < 0.001$ ). Here, higher performance is denoted by a more elevated trendline, delivering faster performance and forecasting better noise tolerance.

### 3.3 Empirical Mosaicking

Based on these results, the multichannel algorithm and optimization of channel selection should enable greater accuracy for real-data mosaicking. Section 2.3 outlines the generation of an empirical data set (without the privilege of ground truth) for this purpose. This empirical set consists of a series of images with varying interframe overlap of a primary tumor cell culture stained using four spectrally distinct fluorescent probes and captured by 32-channel microscopy using a precision stage. We followed the previously discussed steps to optimize the choice of channels for video-rate mosaicking (i.e., limited to up to 10 channels to retain video-rate computation). We estimated  $\epsilon$  using a histogram [Fig. 5(a)] of correlation values ( $R_{ij}$ ) from the first video frame, finding that channels 1, 2, 12, 13, 21, 22, 26, 27, 28, and 32 maximized  $D$ . The images were analyzed by examining the unmixed composite images created by unmixing the



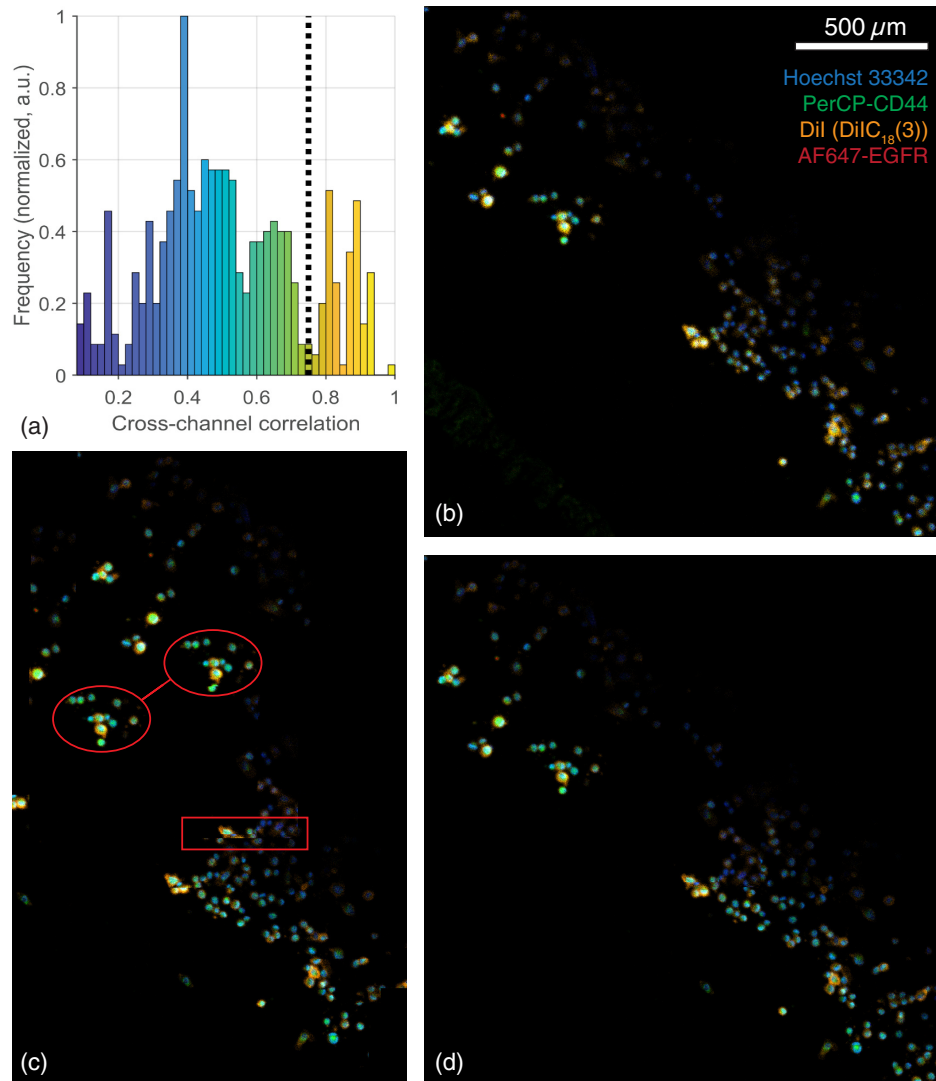
**Fig. 4** Quantification and characterization of mosaicking accuracy, dimensionality score, and computational speed under noisy imaging conditions. (a) Mosaicking accuracy under noise load is modeled using a logistic curve [Eq. (2)] where the logistic midpoint (dashed line) characterizes the noise tolerance. Performance is improved both by adding more channels and optimizing (opt.) the subset of channels utilized versus user-selected channels (ctrl.) based on even sampling of molecular constituents. (b)  $\sigma_{50}^2$  values extracted from midpoints of (a) using Eq. (1) display improved noise tolerance through selection of a channel subset that maximizes  $D$ ; asterisks represent levels of significance (\* signifies  $p < 0.05$ , \*\*\*\* signifies  $p < 0.0001$ ). (c)  $D$  correlates with  $\sigma_{50}^2$  for optimized permutations ( $R^2 = 0.95$ ); data labels denote number of channels used. (d) Mosaicking speed decreases with increasing dimensionality score as a power law. Predicted performance of optimized channel sets is better than that of the user-selected control sets, and unmixed channels provide the best predicted performance. Dotted line denotes video-rate (15 fps) threshold.

raw data into the four endmembers (dyes) and creating a color map to represent all four dyes in one image. The Zeiss tiling mentioned earlier [Fig. 5(b)] provides a gold standard for qualitative comparisons of mosaicking accuracy and for evaluating the presence of artifacts.

The single channel method produced an image [Fig. 5(c)] with many severe artifacts, including sharp edge artifacts and clear examples of feature duplication. These errors are presumed to result from alignment mistakes caused by noise distorting similar areas between frame pairs, ultimately hindering the NCC alignment step. The 10-channel optimized mosaicking [Fig. 5(d)] qualitatively outperformed the one-channel control, producing an image largely free of artifacts, substantially mimicking the gold standard. Additionally, step-by-step comparisons of the video feed and the associated intermediate mosaic can be visualized in Video 1.

This empirical test demonstrates that significant improvement was achieved in a real-data scenario compared to the single-channel analog. Cross-correlation-based algorithms have been implemented successfully for handheld microendoscopes in the literature,<sup>5</sup> so the methods here should achieve similar results for multichannel versions of these probes. However, the primary limitation of the cross-correlation-based method implemented here is that alignments are limited to translational corrections, and therefore nontranslational motion must be negligible. To overcome this limitation, micromosaicking algorithms have been developed far beyond single-channel correlations and now typically utilize feature-based alignment strategies that allow for more complex frame-pair alignments such as affine transformations<sup>7</sup> and local warping corrections.<sup>6</sup> The success of the multichannel NCC method here motivates integration of these advanced methods with the present





**Fig. 5** Micromosaicking of an empirical hyperspectral data set (described in Sec. 2.3). (a) Histogram of correlation values for 32 hyperspectral channels listed in Table S1 in the [Supplemental Materials](#) reveals that a choice of  $\epsilon = 0.75$  (dotted line) separates a population of strongly correlated channels, informing calculation of  $D$  [Eq. (4)]. (b) Tiling of four unmixed dye channels (annotated on image) serves as a qualitative gold standard to compare mosaics. (c) One-channel mosaicking produces severe artifacts such as repeated features (red circle), sharp edges (red rectangle), and a notably different aspect ratio compared to (b) despite being displayed at the same scale. These are indicative of significant misalignments and this configuration ultimately fails to provide an accurate mosaic. (d) Micromosaicking using the optimal 10 channels results in an image free of noticeable artifacts; qualitative comparison to the gold standard (b) indicates that the images have only minute differences (Video 1, 4.93 MB, MP4 [URL: <https://doi.org/10.1117/1.JBO.24.12.126002.1>]).

multichannel mosaicking algorithm in a manner that similarly takes advantage of the unique information available in multispectral and hyperspectral data. A limitation of this method is that smooth image layering such as graph-cut-based stitching<sup>7,24</sup> or alpha blending of overlapping areas<sup>4</sup> have not yet been integrated. The present algorithm also does not compensate for compounding errors (or drift error), where small misalignments may build up due to the pairwise image alignment method and cause a significant discrepancy after many alignments, as it cannot correct itself after misalignments, as demonstrated in Fig. S2(b) in the [Supplemental Materials](#). This problem has been addressed in earlier publications<sup>6</sup> by referencing a global alignment system and refining pairwise alignments to fix these

issues. Here, the results motivate the integration of such methods with the present algorithm in future work.

## 4 Conclusion

In summary, we report a multichannel micromosaicking algorithm capable of processing up to 10 channels (spectral or unmixed molecular basis channels) at video rate (15 fps). This development is motivated by the anticipated need for real-time analysis of multiplexed microscopy and microendoscopy image cube data streams, as hyperspectral technologies continue to advance in speed and spectral resolution. We characterized the noise tolerance of a given grouping of synthetic spectral channels with a dimensionality score that captures two

forms of noise tolerance enforced by the algorithm, and show that this dimensionality score is a reliable predictor of algorithm accuracy and speed performance. This enables the determination of a high-value subset of channels for computational bandwidth-limited situations. The predicted trends from the dimensionality score are matched by actual increases in performance in noisy conditions. We empirically demonstrate the improved performance with a second data set, showing higher mosaic fidelity and fewer artifacts than the single-channel method when compared to an accepted gold standard. We also demonstrate that unmixed endmember abundance maps offer a better optimization of mosaicking speed and dimensionality score than raw spectral channels. This result motivates the further development of accelerated real-time unmixing and its integration with multichannel micromosaicking. Future work includes development of a multichannel feature-based algorithm to unite hyperspectral micromosaicking methods with complex frame pair alignment strategies needed for *in vivo* microscopy using hand-held probes.

### Disclosures

The authors declare no conflicts of interest.

### Acknowledgments

We thank Dr. E. Cronin-Furman (Olympus Corporation of the Americas Scientific Solutions Group) for assistance in acquiring images described in Sec. 2.2 and shown in Fig. S1 in the [Supplemental Materials](#). We also would like to thank the Institute for Chemical Imaging of Living Systems at Northeastern University for assistance in acquiring the images described in Sec. 2.3 and shown in Fig. 5. This work was supported by the U.S. National Institutes of Health Grant No. K22CA181611 (to B.Q.S.) and the Richard and Susan Smith Family Foundation (Newton, Massachusetts) Smith Family Award for Excellence in Biomedical Research (to B.Q.S.).

### References

1. B. A. Flusberg et al., "Fiber-optic fluorescence imaging," *Nat. Methods* **2**(12), 941–950 (2005).
2. B. W. Pogue et al., "Perspective review of what is needed for molecular-specific fluorescence-guided surgery," *J. Biomed. Opt.* **23**(10), 100601 (2018).
3. M. Irani, S. Hsu, and P. Anandan, "Video compression using mosaic representations," *Signal Process. Image Commun.* **7**(4–6), 529–552 (1995).
4. T. Vercauteren et al., "Robust mosaicking with correction of motion distortions and tissue deformations for *in vivo* fibered microscopy," *Med. Image Anal.* **10**(5), 673–692 (2006).
5. N. Bedard et al., "Real-time video mosaicking with a high-resolution microendoscope," *Biomed. Opt. Express* **3**(10), 2428–2435 (2012).
6. K. E. Loewke et al., "In vivo micro-image mosaicking," *IEEE Trans. Biomed. Eng.* **58**(1), 159–171 (2011).
7. K. Kose et al., "Automated video-mosaicking approach for confocal microscopic imaging *in vivo*: an approach to address challenges in imaging living tissue and extend field of view," *Sci. Rep.* **7**(1), 10759 (2017).
8. R. A. Schultz et al., "Hyperspectral imaging: a novel approach for microscopic analysis," *Cytometry* **43**(4), 239–247 (2001).
9. N. Hagen and M. W. Kudenov, "Review of snapshot spectral imaging technologies," *Opt. Eng.* **52**(9), 090901 (2013).
10. G. Lu and B. Fei, "Medical hyperspectral imaging: a review," *J. Biomed. Opt.* **19**(1), 010901 (2014).
11. H. Makhoulf et al., "Multispectral confocal microendoscope for *in vivo* and *in situ* imaging," *J. Biomed. Opt.* **13**(4), 044016 (2008).
12. R. T. Kester et al., "Real-time snapshot hyperspectral imaging endoscope," *J. Biomed. Opt.* **16**(5), 056005 (2011).
13. C. L. Zavaleta et al., "A Raman-based endoscopic strategy for multiplexed molecular imaging," *Proc. Natl. Acad. Sci. U. S. A.* **110**(25), E2288–E2297 (2013).
14. D. Cerra, R. Müller, and P. Reinartz, "Noise reduction in hyperspectral images through spectral unmixing," *IEEE Geosci. Remote Sens. Lett.* **11**(1), 109–113 (2014).
15. S. K. Misra et al., "Hyperspectral imaging offers visual and quantitative evidence of drug release from zwitterionic-phospholipid-nanocarbon when concurrently tracked in 3-D intracellular space," *Adv. Funct. Mater.* **26**(44), 8031–8041 (2016).
16. R. Miranda-Luna et al., "Mosaicking of bladder endoscopic image sequences: distortion calibration and registration algorithm," *IEEE Trans. Biomed. Eng.* **55**(2), 541–553 (2008).
17. R. M. Willett et al., "Sparsity and structure in hyperspectral imaging: sensing, reconstruction, and target detection," *IEEE Signal Process. Mag.* **31**(1), 116–126 (2014).
18. B. Q. Spring and R. M. Clegg, "Image analysis for denoising full-field frequency-domain fluorescence lifetime images," *J. Microsc.* **235**(2), 221–237 (2009).
19. J. Nickolls and W. J. Dally, "The GPU computing era," *IEEE Micro* **30**(2), 56–69 (2010).
20. S. Bernabé et al., "Multicore real-time implementation of a full hyperspectral unmixing chain," *IEEE Geosci. Remote Sens. Lett.* **15**(5), 744–748 (2018).
21. R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*, Vol. II, Addison-Wesley, Reading, Massachusetts (1992).
22. D. W. Hosmer and S. Lemeshow, *Applied Logistic Regression*, 2nd ed., Wiley, New York (2000).
23. S. Dowdy and S. Wearden, *Statistics for Research*, Wiley, New York (1983).
24. Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(11), 1222–1239 (2001).

Biographies of the authors are not available.