Journal of Medical Imaging

MedicalImaging.SPIEDigitalLibrary.org

Application of unsupervised learning to hyperspectral imaging of cardiac ablation lesions

Shuyue Guan
Huda Asfour
Narine Sarvazyan
Murray Loew

SPIE.

Shuyue Guan, Huda Asfour, Narine Sarvazyan, Murray Loew, "Application of unsupervised learning to hyperspectral imaging of cardiac ablation lesions," *J. Med. Imag.* **5**(4), 046003 (2018), doi: 10.1117/1.JMI.5.4.046003.

Application of unsupervised learning to hyperspectral imaging of cardiac ablation lesions

Shuyue Guan,[a] Huda Asfour,[b] Narine Sarvazyan,[b] and Murray Loew[a,*]
[a]George Washington University, Department of Biomedical Engineering, Washington, DC, United States
[b]George Washington University Medical Center, Department of Pharmacology and Physiology, Washington, DC, United States

Abstract. Atrial fibrillation is the most common cardiac arrhythmia. It is being effectively treated using the radiofrequency ablation (RFA) procedure, which destroys culprit tissue and creates scars that prevent the spread of abnormal electrical activity. Long-term success of RFA could be improved further if ablation lesions can be directly visualized during the surgery. We have shown that autofluorescence-based hyperspectral imaging (aHSI) can help to identify lesions based on spectral unmixing. We show that use of *k*-means clustering, an unsupervised learning method, is capable of detecting RFA lesions without *a priori* knowledge of the lesions' spectral characteristics. We also show that the number of spectral bands required for successful lesion identification can be significantly reduced, enabling the use of increased spectral bandwidth. Together, these findings can help with clinical implementation of a percutaneous aHSI catheter, since by reducing the number of spectral bands one can reduce hypercube acquisition and processing times, and by increasing the spectral width of individual bands one can collect more photons. The latter is of critical importance in low-light applications such as intracardiac aHSI. The ultimate goal of our studies is to help improve clinical outcomes for atrial fibrillation patients.
[DOI: 10.1117/1.JMI.5.4.046003]

Keywords: autofluorescence; cardiac ablation; hyperspectral imaging; radiofrequency ablation; unsupervised learning; lesion detection/visualization; *k*-means clustering; spectral unmixing.

Paper 18107R received May 22, 2018; accepted for publication Nov. 12, 2018; published online Dec. 15, 2018.

1 Introduction

Atrial fibrillation (AF) is the most common cardiac arrhythmia, affecting as many as 10 million people in the United States alone.[1] Most of the abnormal sources of bioelectrical activity causing AF have been found in the left atrium near the entrance of the pulmonary veins.[2] AF can be treated by destroying the culprit tissue, thereby creating scar tissue, which prevents abnormal activity from spreading.[3] Radiofrequency ablation (RFA) is a common surgical procedure being used widely to ablate living tissues, including in the atria. Testing of electrical conduction is then used to determine if abnormal sources of electrical activity have been isolated. Even after passing such testing, however, patients can later experience repeat AF arising because reversible tissue injury and temporary edema can also stop electrical activity.[4] When tissue recovers, electrical reconnections can lead to AF recurrence. The recurrence rate of AF after an ablation procedure can be as high as 50% and more than 90% of these recurrent cases have been linked to gaps between ablation lesions.[5–7] Incomplete placement of lesions that later result in AF recurrence can be curtailed if clinicians could directly monitor lesion formation along with the degree of tissue damage. Unfortunately, the endocardial surface of the left atria, where most of RF ablation procedures are performed, is covered by thick layers of collagen and elastin preventing direct visualization of ablated muscle beneath. While imaging technologies such as MRI, CT, and ultrasound have been successfully applied for lesion testing, they have significant limitations. CT and MRI are expensive, involve radiation and/or contrast agents, and ultrasound imaging has poor image resolution.[8,9] Therefore, our group has been exploring a visualization approach called autofluorescence hyperspectral imaging (aHSI) and has shown that it is effective in revealing ablation-induced damage including the highly collagenous human left atrium.[10–14]

To implement the aHSI approach during the RFA procedure, one has to deliver ultraviolet (UV) light ($\lambda = 365$ nm) to the heart by an optical fiber threaded into a percutaneous catheter.[15,16] This allows illumination of the endocardial atrial surface, which is highly autofluorescent. The autofluorescence signal is then detected through the image guide and the attached HSI camera system, which forms a stack of images acquired at individual wavelengths. Figure 2 shows a summary diagram of such a system, and Fig. 3 shows the hypercube construction. The hypercubes contain rich spectral information about the tissue. Our previous studies have shown that subtle changes in the tissue autofluorescence profiles can help to identify the ablated regions in both animal and human atrial tissues.[10,11] In those studies, we had to preacquire target spectra for lesion and nonlesion sites before applying linear unmixing,[10] since it is a supervised learning method. The first objective of this work was to apply an unsupervised learning method, *k*-means clustering, to detect RFA atrial lesions without *a priori* knowledge about tissue spectra. Our second objective was to use *k*-means clustering to select the minimal number of spectral bands (feature groups)

*Address all correspondence to: Murray Loew, E-mail: loew@gwu.edu

2329-4302/2018/$25.00 © 2018 SPIE

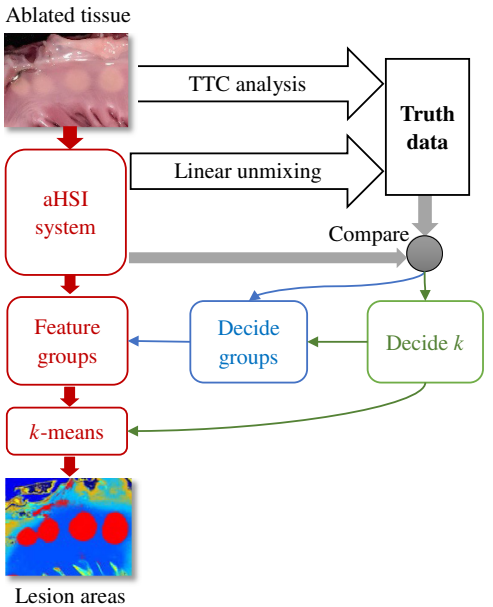

Fig. 1 The flowchart summarizes the methods used in this study.

without significantly reducing the accuracy of lesion detection. This is important for future implementation of an intracardiac aHSI catheter, since it is beneficial to decrease the number of spectral images within the hypercube while preserving the method's ability to reveal the lesions. First, having fewer images will speed-up both acquisition and processing, enabling us to visualize the ablated areas in real time. Second, by widening spectral bands around the most useful wavelengths, one can collect more photons and make the output images more robust to noise.

Figure 1 briefly summarizes this study. The ground truth areas of lesion were obtained from the aHSI system by linear unmixing and verified by 2,3,5-triphenyl-2H-tetrazolium chloride (TTC) analysis. By comparing with truth data, we found the optimal k-value for k-means algorithm (green) as well as the optimal groups (blue). The procedure on the left (red) is our proposed methods for lesion detection from ablated tissue to lesion areas. More details about the procedures of this study are shown in the flowchart in Sec. 6.

2 Materials and Methods

2.1 *Hyperspectral Imaging Hardware*

Atria from freshly excised porcine hearts were ablated by a nonirrigated RF ablation catheter (Boston Scientific). Several lesions were created on one tissue sample. Atria were illuminated with a 365-nm UVA LED (Mightex, Pleasanton, California) placed ~10 cm from the tissue surface. A CCD camera outfitted with a Nikon AF Micro-Nikkor 60 mm $f/2.8D$ objective and a liquid crystal tunable filter (LCTF, Nuance FX, PerkinElmer/CRi) was used to acquire hypercubes of the samples.

The LCTF was tuned to pass the wavelengths from 420 to 720 nm at wavelengths separated by the filter's band interval, 10 nm; this yielded 31 channels. We note that the full-width at half-maximum (FWHM) of the response at each wavelength is specified as a constant (20 nm) for the Nuance FX.[17] The effect of the FWHM on spectral resolution is minimal, in light of the subsequent grouping of bands (see Sec. 2.3.2). As shown in Fig. 2, through the LCTF, a lens projects the collected light onto a CCD containing 1392×1040 pixels. Finally, the hypercube for each sample was constructed from the 31 autofluorescence images, each of size 1392×1040 (see Fig. 3). Ten samples were used in this study; therefore, we collected 310 autofluorescence images in total.

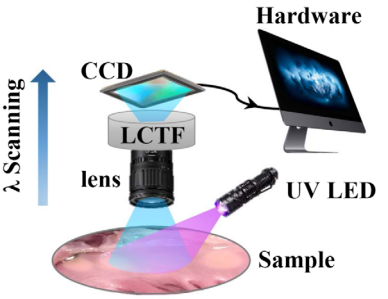

Fig. 2 Schematic showing hypercube acquisition. CCD, charge coupled device; LCTF, liquid crystal tunable filter; UV LED, ultraviolet light-emitting diode.

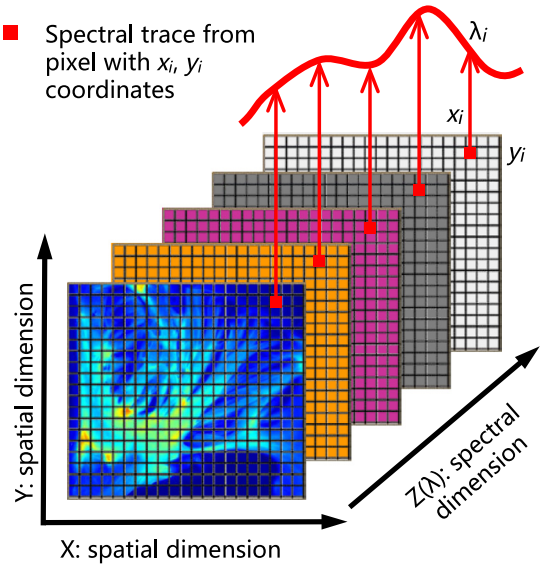

Fig. 3 Hypercube of aHSI images: images in the hypercube were ordered by their wavelength increasingly on the Z-axis. Each pixel on the $X - Y$ plane thus has an associated spectrum.

2.2 *Data Preprocessing*

For each sample, we combined the 31 images into a three-dimensional (3-D) hypercube and extracted spectral profiles from each x and y pixel. Each spectrum was then divided by the spectral sensitivity curves of the CCD camera and the LCTF[10] (correction), followed by normalization, which converted values of each spectrum to the range from 0 to 1 [Fig. 4(a)]. Normalization is critical because for the classification algorithm it is the overall shape of the spectrum that matters, but not the absolute light intensity at each wavelength. For normalization, the maximum value was set at 1 and the minimum value at 0. More details about the importance of normalization step are included in the reference to our earlier work.[10]

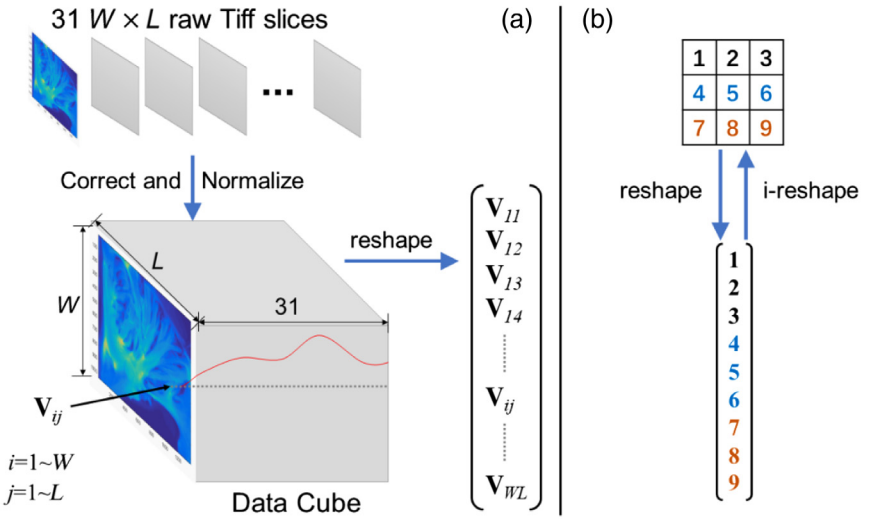

Fig. 4 (a) Preprocessing operations and reshaping hypercube into a 2-D matrix and (b) the rule of reshaping and inverse-reshaping.

Then, we reshaped the 3-D hypercube to a two-dimensional (2-D) matrix according to the rule shown in Fig. 4(b): for every point on the $X-Y$ plane (a pixel), the data along the spectral dimension were considered as a vector in the new 2-D matrix; the pixels were ordered from left to right in the first row (upper left), then in the second row and so forth. The spectrum of each pixel in the $X-Y$ plane was represented as a vector in the matrix; the matrix therefore had 31 columns corresponding to 31 spectral bands (420 to 430 nm, 430 to 440 nm,...,710 to 720 nm). Hereafter, we refer to each pixel as a sample; each sample is a vector of 31 features.

2.3 *Unsupervised Learning and Lesion Detection*

2.3.1 *Background*

Unsupervised learning methods can infer the hidden structures or extract information from unlabeled data.[18,19] Its advantage is that *a priori* knowledge (e.g., labels) about the targets is not required before performing detection. Unsupervised learning algorithms have been applied to target detection in various fields, such as anomaly detection,[20] road detection,[21] object recognition,[22] and salience detection.[23–26]

In this study, we used k-means clustering[27] as an unsupervised learning algorithm to cluster samples (vectors of spectra) into k clusters, numbered by integers from "1" to "k". Each location (pixel) of the spectrum was labeled with its cluster number. Then, we assigned colors to those numbers to allow visualization of the clusters. Since the spectra of lesions and nonlesions are different, they were assigned different colors to be distinguished visually from the other tissues.

The clustering algorithm was performed in a computer with Intel® Core™ i7-6700 3.40 GHz CPU and 16.0 GB RAM. Its operating system is Windows 10 and application is MATLAB R2016b. The built-in k-means function in MATLAB applies the squared Euclidean distance measure and the "k-means++" algorithm for initializing cluster centers; and its maximum number of iterations is 100.

2.3.2 *Grouping and clustering*

Our goal was to use band grouping to decrease the number of spectral components to at most four. If one can use four composite spectral bands, this will enable four quadrants of the camera chip to be used simultaneously (see Sec. 2.6).[28] This also has the benefit of increasing the signal-to-noise ratio, which is essential because of the weak fluorescence and its attenuation by the optical fiber. The next step was to use machine learning algorithms to detect lesion areas through those four features. The effectiveness of a given set of four groups was assessed using k-means clustering.[27] In our case, we have $1040 \times 1392 = 1{,}447{,}680$ samples for each dataset. We performed k-means clustering, in which the value of k is unknown initially and determined by experiment. Each pixel was labeled by its cluster. Then, we assigned colors to these numbers to allow visualization of the clusters. The procedure is shown in Fig. 5.

The k-means clustering method is a commonly used unsupervised machine learning algorithm. In general, given an input set $\{\mathbf{x}_i\}$ having m d-dimensional real vectors, we use k-means clustering to partition the m vectors into $k(\leq m)$ sets $\mathbf{S} = \{s_1, s_2, \ldots, s_k\}$ and minimize the within-cluster sum of squares

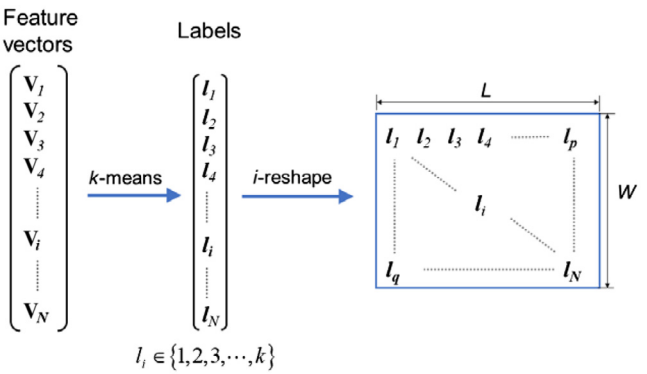

Fig. 5 K-means clustering.

$$\arg \min_{s} \sum_{j=1}^{k} \sum_{\mathbf{x} \in s_j} \|\mathbf{x} - \boldsymbol{\mu}_j\|^2,$$

where $\boldsymbol{\mu}_j$ is the mean of vectors in s_j. It is implemented with Lloyd's algorithm.[29]

2.4 *Creating Reference Images*

To be able to evaluate any lesion detection method, one must have sets of images in which the lesions are labeled. This section describes the construction of such sets. A traditional way to outline the lesions is to stain tissue with TTC. The content of dehydrogenase enzymes and NADH declines within ablated tissue. Since those compounds turn tetrazolium salts into a formazan pigment, viable tissue turns red, while lesion areas appear white [Fig. 6(a)]. TTC staining thus provides reliable identification of lesions and their boundaries; this is TTC-reference.

Our previous studies have shown that lesions detected by the linear unmixing algorithm based on preacquired spectral libraries closely match RF lesions in the corresponding TTC image (TTC-reference).[10,11] Figure 6(b) shows one such example. Therefore, we considered the lesion component image obtained using linear unmixing of a 31-band hypercube as a reference image; this is called gray-unmixing-reference.

The gray-unmixing-reference has a continuous gray-scale, and lesions are brighter (have larger gray values) than nonlesion areas. To create a new image that identifies unambiguously the lesion and nonlesion pixels, we used a gray-level threshold. The threshold was found by Otsu's method,[30] which uses the image's histogram to find the threshold that maximizes the between-class variance. The pixels with intensities greater than the threshold then were then labeled "lesion;" all others were considered "nonlesion." Having this binary (two-class) image (the biunmixing-reference), enabled us to then quantitively evaluate the k-means approach.

K-means clustering yielded an image in which each pixel was labeled with an integer from "1" to "k." For finding the label of lesions, we recorded the locations of all lesion pixels (whose value is "1") in the biunmixing-reference. Then, we examined all the corresponding pixels (those having the same locations) in the k-means image. Since every pixel has a label (cluster number) after k-means clustering, we can calculate the modal (most-often occurring) label of these sample pixels as the label of lesions; all other labels represent nonlesions. Finally, in the clustering image, all pixels having the label of lesion were set to value "1;" and other pixels (nonlesion) were set to value "0." So, we obtained the binary image (the bi-31-result) for lesion detection by k-means clustering using 31 features.

2.5 *Evaluation by Accuracy Index*

To verify whether k-means clustering using 31 features is an effective method to detect lesion areas, we measured pixel-to-pixel matching by comparing the bi-31-result [example: lesions colored red in Fig. 8(d)] with the outcome of linear unmixing [lesion areas in bi-unmixing-reference; example: white regions in Fig. 8(f)].

The bi-unmixing-reference (I_{Ref}) and bi-31-result ($I_{31\text{Rlt}}$) are binary images having the same size; if the value of a given pixel was different in the two images, it was declared to be a "miss." Accuracy index (Acc) was defined as 1 minus the ratio of the number of "miss" (Diff) to the total number (N) of pixels of lesion areas in the two (detected and truth) images:

$$\text{Acc}(I_{\text{Ref}}, I_{31\text{Rlt}}) = 1 - \frac{\text{Diff}}{N}.$$

If the accuracy was acceptable, we could use the lesions that were detected by k-means using 31 features as a reference (bi-31-result) to evaluate the outcomes after the next step: feature grouping.

2.6 *Feature Grouping*

We implemented a grouping procedure, which divides the 31 features into four contiguous disjoint groups. For each group, we calculated the sum of values as a new feature value, yielding four new features (see Fig. 7).

There are 4060 ways to divide 31 features into four separate and contiguous groups. The 31 features are the intensities at each of the wavelengths from 420 to 720 nm. The goal was to find the best four-feature groupings from the 4060 possible combinations to adequately detect the lesion areas. That number is sufficiently small that we could construct every possible grouping and get its detection result (the bi-4-result).

We assigned a serial number (SN) to each combination. The boundaries between groups were described by the last feature's number in the first, second, and third group. "720" is not shown because it is always the last feature's number in the fourth group (Table 1).

We assess the bi-4-results ($I_{4\text{Rlt}}$) by comparing them to the bi-31-results ($I_{31\text{Rlt}}$) to yield the accuracy: $\text{Acc}(I_{31\text{Rlt}}, I_{4\text{Rlt}})$ whose calculation method was the same as $\text{Acc}(I_{\text{Ref}}, I_{31\text{Rlt}})$.

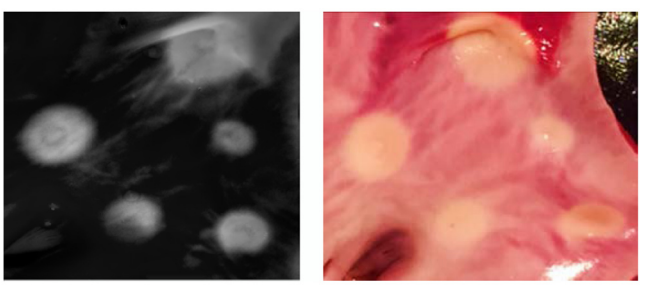

Fig. 6 Appearance of ablated tissue after: (a) linear unmixing from aHSI system and (b) TTC staining.

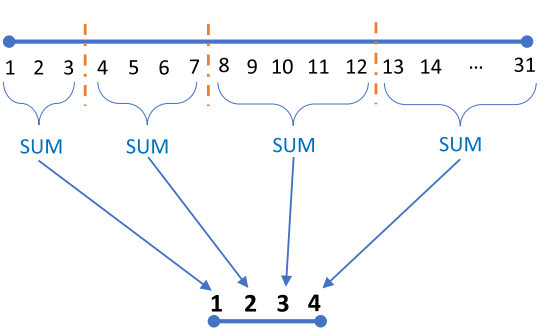

Fig. 7 One kind of four-feature grouping.

Table 1 Combinations of four groups.

SN	Three dividers for four-feature grouping (nm)
1	420, 430, 440
2	420, 430, 450
3	420, 430, 460
...	...
28	420, 430, 720
29	420, 440, 450
30	420, 440, 460
...	...
4060	690, 700, 710

3 Results

3.1 *Lesion Detection by k-Means Clustering*

For porcine samples that encompassed an area of 1392×1040 pixels, Fig. 8(a) shows that $k = 5$ is not sufficient to distinguish ablated regions for this sample (Set-1). To find the optimal k, we computed k ranging from 2 to 41 for all our porcine datasets. For each k, we plotted the maximum, average, and minimum accuracies over the 10 datasets in Fig. 9.

Because a smaller k will make k-means run faster, we seek the smallest k that is effective. Figure 9 shows that $k = 10$ is overall optimal: it is the smallest k that almost reaches all the highest values for maximum, average, and minimum accuracies. As shown in Fig. 8(b), $k = 10$ is effective for the Set-1 sample.

A set of 31 aHSI planes was required to obtain the lesion detection results shown above. The accuracy of lesion detection

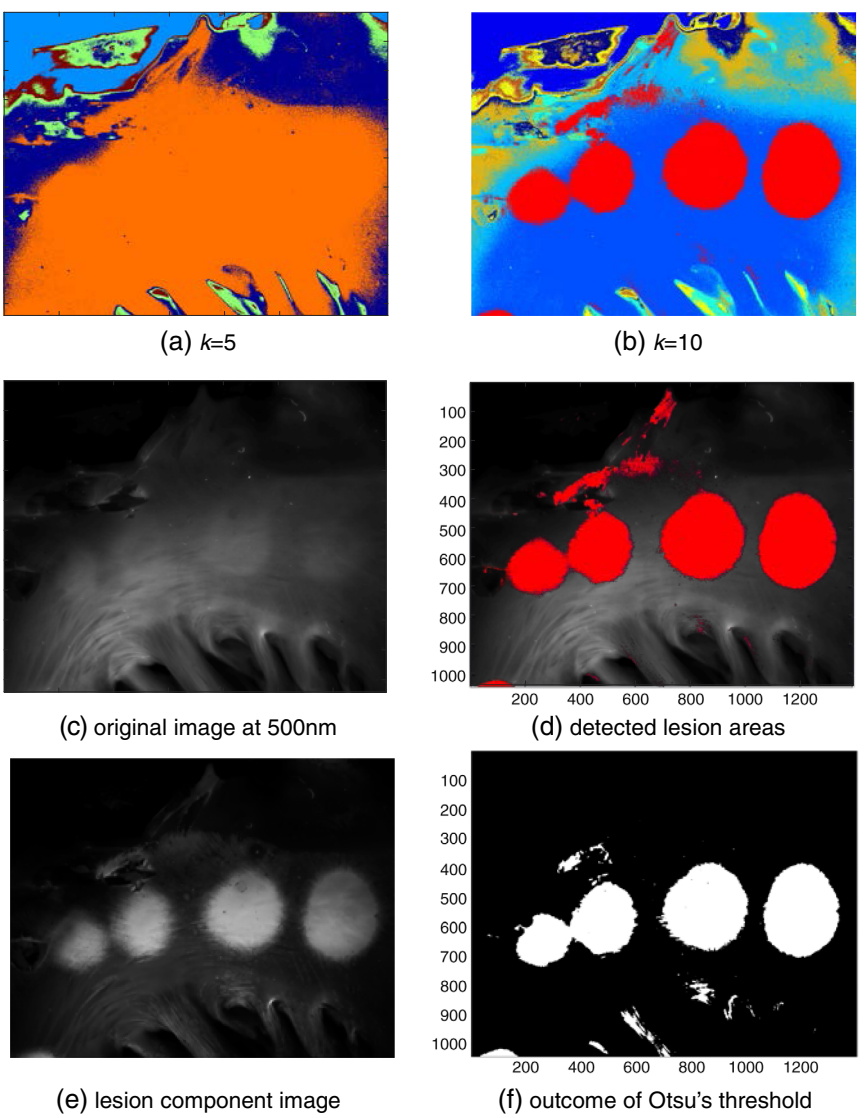

Fig. 8 Results for porcine atria (set-1) clustered by k-means into: (a) five clusters and (b) 10 clusters. (c) An autofluorescence image at 500 nm; (d) the lesion areas (red) detected when $k = 10$, superimposed on the image in (c). The corresponding lesion component image, which is from the unmixed image that contains lesion and nonlesion components, is shown in (e); followed by binary image obtained from (e) by applying Otsu's thresholding (f).

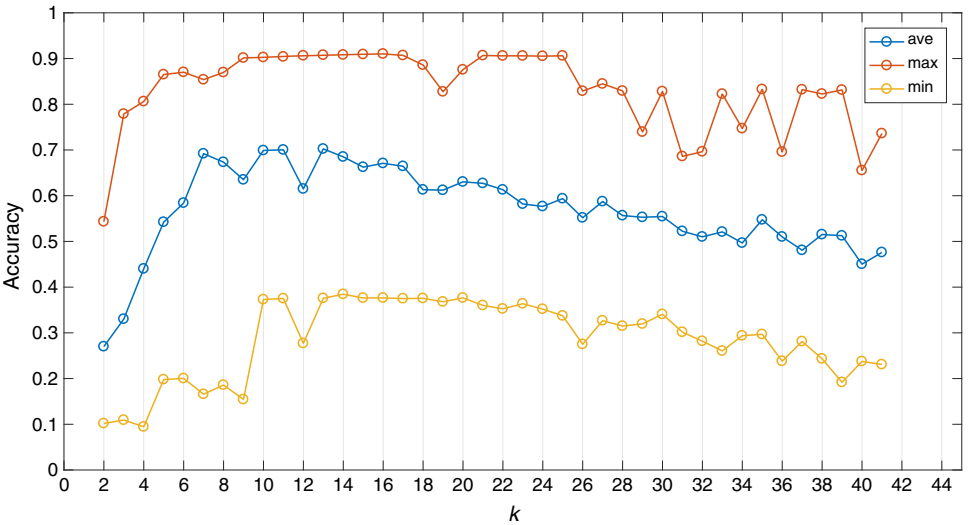

Fig. 9 Maximum, average, and minimum accuracies over 10 datasets for each k.

Table 2 Accuracies of 31-feature clustering results.

Dataset	1	2	3	4	5	6	7	8	9	10
Acc(I_{Ref}, $I_{31\text{Rlt}}$)	0.87	0.91	0.38	0.80	0.69	0.82	0.66	0.87	0.66	0.75

by $k = 10$-means clustering using 31 aHSI planes has been measured by comparison with the images revealed by linear unmixing (bi-unmixing-reference). Table 2 shows their accuracies. Note that dataset 3 has lower accuracy than the others because Otsu's algorithm cannot automatically extract a good binary reference image on this dataset.

The average accuracy for detection by $k = 10$-means using 31 features was about 74% when evaluated using the bi-unmixing-reference. We then grouped the features. The goal was to decrease the number of spectral bands without reducing the accuracy of lesion detection appreciably.

3.2 *Feature Grouping*

Using the same sample, by comparing its bi-4-result with bi-31-result pixel by pixel, we calculated the accuracy of the feature grouping. In this experiment, we computed the accuracies for 4060 combinations of four groups for one sample (set-1); the accuracy for each SN is shown in Fig. 10. (The periodicity apparent in Figs. 10–13 is an artifact of the serial-numbering system, and has no significance.)

By analyzing Fig. 10, we found the feature grouping having the highest accuracy of lesion detection for this specific sample.

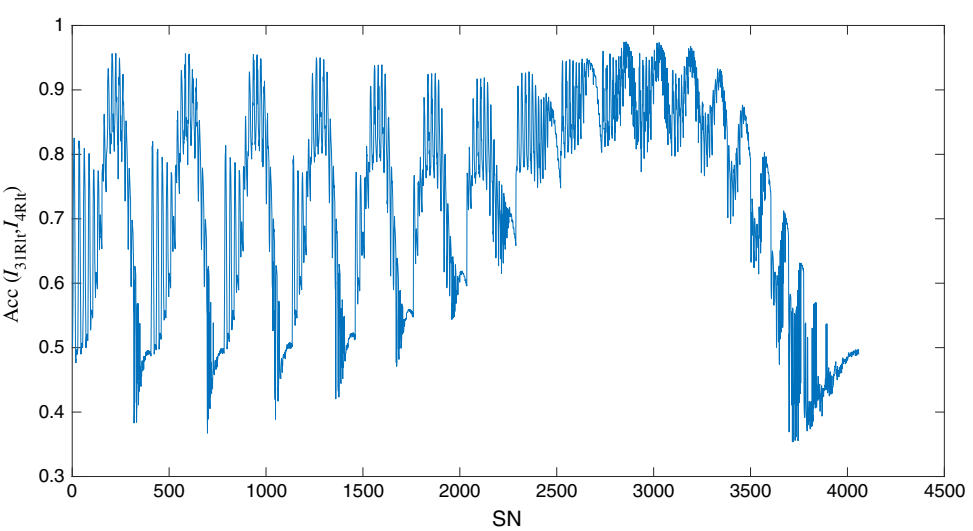

Fig. 10 Accuracies of SNs for one dataset (set-1).

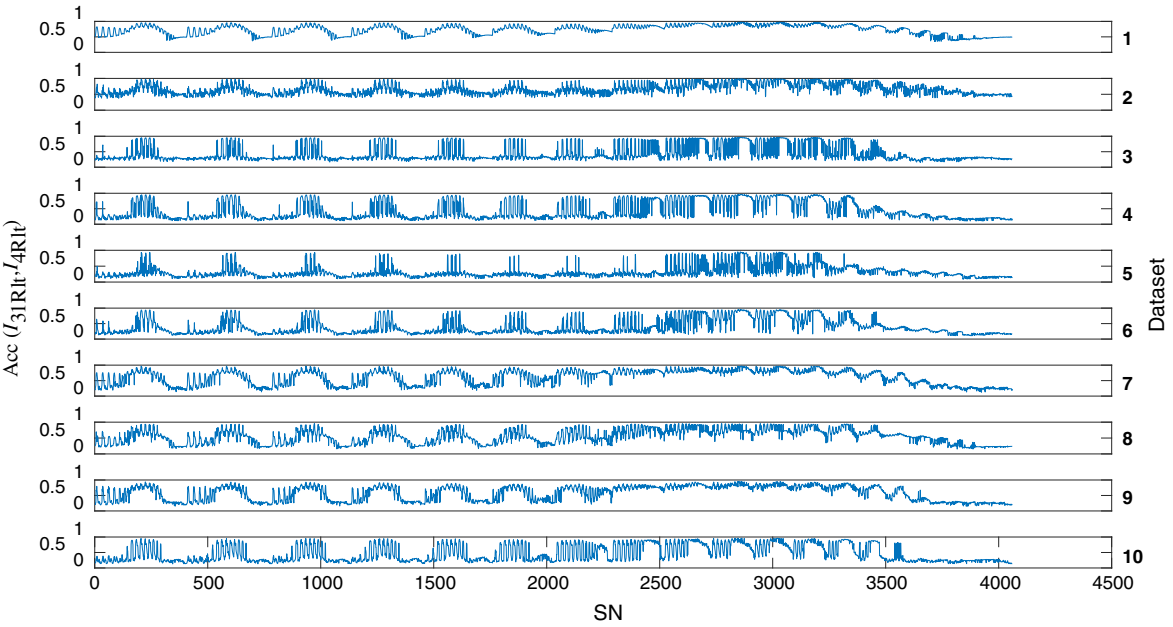

Fig. 11 Feature grouping accuracies for 10 datasets; each row represents a dataset.

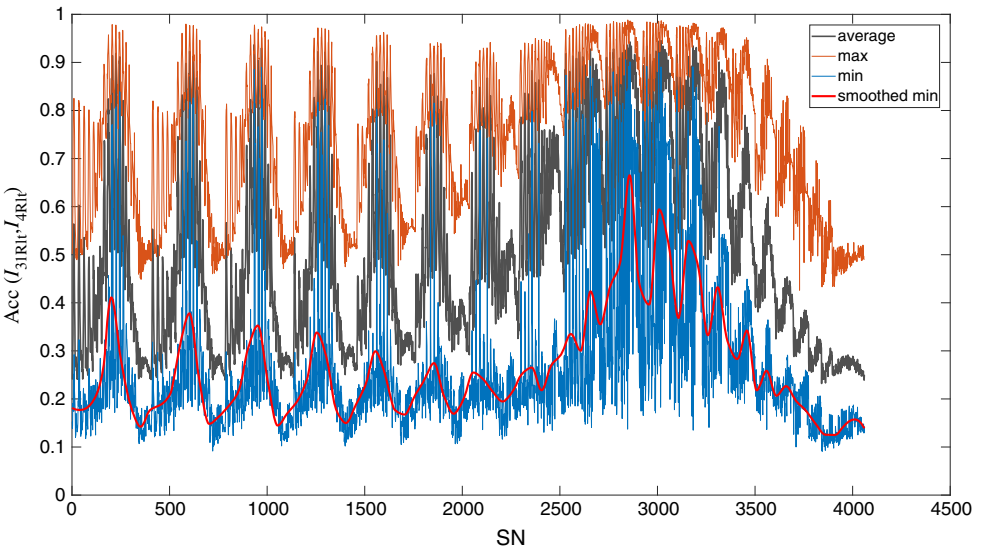

Fig. 12 Accuracies over 10 datasets.

As shown in Fig. 14, the four-feature result with higher accuracy (b) identifies more complete lesion areas than the 4-feature result with lower accuracy (c).

We then applied the same approach to the remaining datasets. The goal was to find four-feature grouping combinations that have high accuracies of lesion detection across multiple studies. To do this we examined the accuracies for all 10 datasets (Fig. 11). As Fig. 11 shows, the accuracy of a given feature grouping varies across datasets. One group-selection method is to record the worst performance of a feature grouping across all datasets:

$$W_{\mathrm{SN}} = \min_i \{A^i_{\mathrm{SN}}\},$$

where A^i_{SN} is the accuracy of the i'th dataset and SN is the serial number.

The result (Fig. 12) shows that there are many four-feature groupings that perform well (peaks) across all samples we tested. But taking the worst performance has a flaw: if there exists one bad dataset, reducing accuracies of all feature groupings appreciably, it will influence the result greatly. If we plot the smoothed (by average) minimum accuracies, obviously periodic phenomenon to series numbers is shown. Figure 15 shows the locations (wavelength) of dividers and the smoothed minimum accuracy value (scaled) to each SN. And, by looking at this figure, one can notice that such periods are defined by the first divider. We also noticed that the SN range in the green area (2730 to 3245), in which the divider

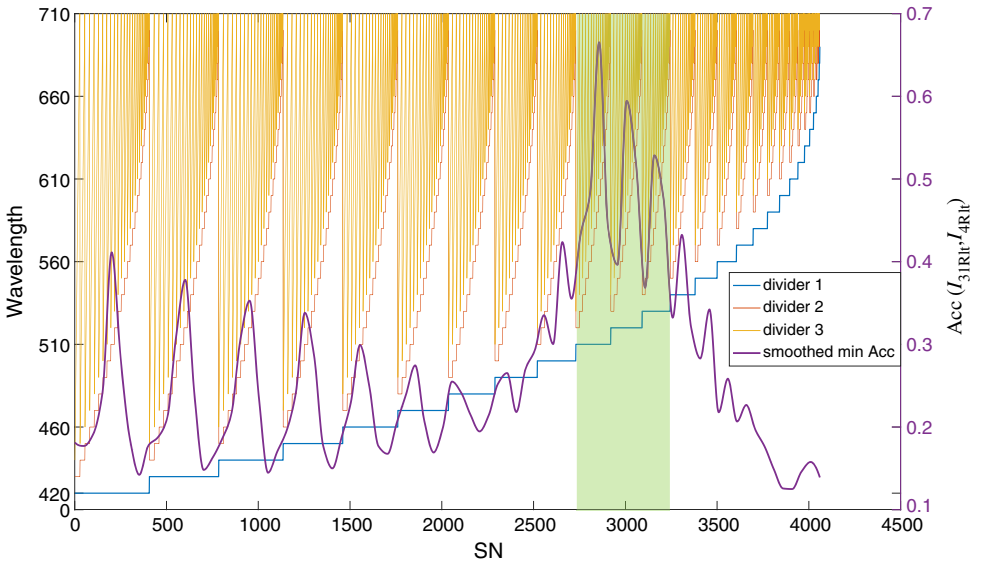

Fig. 13 Smoothed minimum accuracies (scaled) with the three dividers.

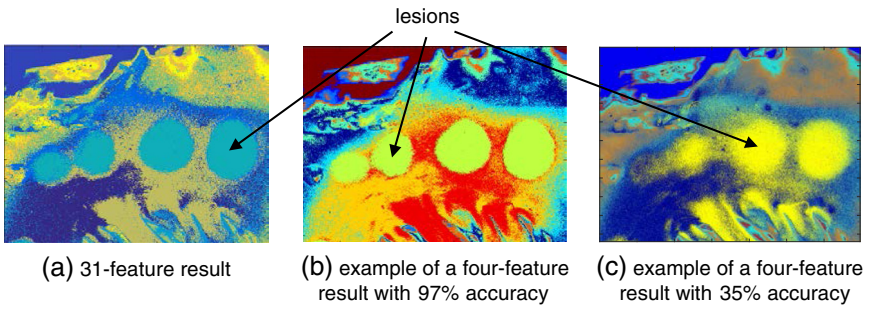

Fig. 14 Feature grouping results for porcine atria (set-1): (a) *k*-means clustering ($k = 10$) by using all 31 features; (b) *k*-means clustering ($k = 10$) by using four features from 4-feature grouping (SN = 2857): (wavelength groups: 420 to 510, 520 to 600, 610 to 630, and 640 to 720 nm); (c) *k*-means clustering ($k = 10$) using four features from different four-feature grouping (SN = 3716): (wavelength groups: 420 to 580, 590 to 600, 610 to 680, and 690 to 720 nm).

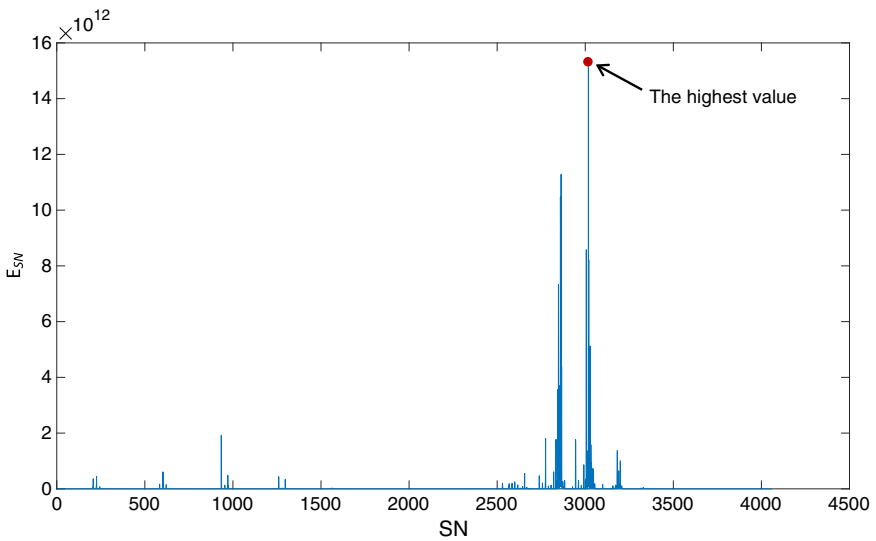

Fig. 15 Evaluated accuracies over 10 datasets.

Table 3 Accuracies of four-feature clustering results by grouping (SN = 3018): (420 to 520, 530 to 590, 600 to 640, and 650 to 720 nm).

Dataset	1	2	3	4	5	6	7	8	9	10
Acc($I_{31\text{Rlt}}, I_{4\text{Rlt}}$)	0.97	0.99	0.96	0.96	0.92	0.93	0.96	0.93	0.89	0.95

1 ranges from 510 to 530 nm, includes most high-accuracy combinations.

Another method is to design an evaluation function to reflect the average performance of a feature grouping combination through all samples:

$$E_{\text{SN}} = \prod_i \frac{1}{1 - A_{\text{SN}}^i},$$

where A_{SN}^i is the accuracy of i'th dataset and SN is the serial number of combination.

By this equation, a feature grouping combination will get a large score if its accuracy is close to 1. Thus, this nonlinear function emphasizes high accuracy values. Since the maximum value of accuracy in this work was less than 0.99, E_{SN} is bounded. By comparing the two results (Figs. 12 and 15), we observe that the good-performance combinations are similar. The evaluation function could find a better grouping for all tested datasets than the minimum accuracy method, but the grouping that we obtained from the max-min method could be more stable for new datasets because it provides a reliable lower bound on accuracy.

From this evaluation, the highest value (red point) is obtained with the four-feature grouping (420 to 520, 530 to 590, 600 to 640, and 650 to 720 nm) (SN = 3018). Using this grouping, we found the accuracies of four-feature clustering results for all datasets (Table 3).

3.3 *Time Cost for k-Means Clustering*

An important practical issue is the computing time cost for k-means clustering based on all 31 features and that based on the four grouped features. We used 10 datasets to test the time cost of k-means clustering. The main factors that affect the computation time are the number of sample vectors, the dimension of vectors (the number of features), maximum iterations of k-means, and the value of k. For one dataset, the number of sample vectors is fixed ($1040 \times 1392 = 1{,}447{,}680$), and the default maximum iterations of k-means in MATLAB is 100. Therefore, the running time depends primarily on the number of features and the value of k. The running times of k-means may vary for each run. For a given k, the time cost of four-feature clustering is about 41.3% of that of the 31-feature clustering, while the average accuracy of four-feature clustering by grouping (420 to 520, 530 to 590, 600 to 640, and 650 to 720 nm) is about 95% of that of the 31-feature clustering (Fig. 16). We conclude that the four-feature grouping can greatly speed up the processing while maintaining good accuracy of lesion detection.

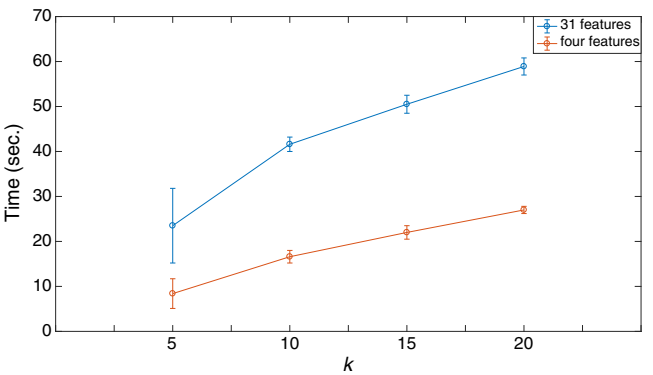

Fig. 16 Time costs for k-means clustering.

4 Discussion

We used the k-means clustering method to find the lesion sites and compared the outcomes to those using linear unmixing. Since k-means is an unsupervised learning algorithm, we did not require *a priori* knowledge of lesion spectra. In contrast, the supervised learning methods do require such knowledge about the lesion to construct the training set containing labeled spectra. In practice, k-means assigns lesion and nonlesion areas to different groups and assigns different colors. The outcome of k-means verified our hypothesis that the spectra of ablated tissue are different from those of nonablated tissue. Also, it confirmed that the autofluorescence images contain information about the components and structure of tissues.[10] Whereas k-means clustering is in general repeatable, one disadvantage is that the detected lesion regions may vary slightly for each clustering result from a given dataset. That is a characteristic of k-means because the initial points of groups are selected randomly, and the clustering result may be affected by the choice of initial points.

Alternatively, we could apply supervised learning methods. A classifier would be trained through labeled lesion and nonlesion spectral data. One advantage of supervised classification is that the regions of lesion detected by a classifier model are invariant for a given dataset. Though the time for training a classifier might be greater, the lesion detection process by using the classifier would be faster unless the classifier model is very complicated (nonlinear and in high dimension). But its disadvantage is that one will require a large amount of labeled lesion and nonlesion spectral data for such training.

In future work, we could use spectral unmixing to group the features instead of k-means clustering. As a supervised detection method, the end-member spectra used for unmixing are generated from labeled lesion and nonlesion spectral data. An advantage of spectral unmixing is that its detected result yields continuous values for the intensities of lesions, from which one may be able to infer the depths and 3-D shapes of lesions.[10,31] In contrast, the detected areas provided by k-means are binary images.

To evaluate the results presented in this paper, we compared the outcomes after feature grouping with the results before feature grouping ($I_{4\text{Rlt}}$ versus $I_{31\text{Rlt}}$). In addition, the results

before feature grouping were verified by comparing them with the outcomes of linear unmixing ($I_{31\text{Rlt}}$ versus I_{Ref}). A direct comparison between the outcomes of k-means and TTC staining would have been ideal, but this presents practical problems. First, the chemical reaction that occurs during TTC staining makes ablated tissues shrink to a certain degree. Second, images taken after TTC staining are not taken at exactly the same orientation, so an exact comparison is not possible, even when image registration methods are used. But since the main goal of this paper was to find the best feature grouping, direct comparison with TTC was not necessary; and we have previously reported a direct comparison between lesion surface areas of the lesions in TTC images and those obtained in gray unmixing images.[10,31] By computing the difference between detected lesions before and after feature grouping, we still were able to achieve our goal.

The best four feature grouping or spectral band areas corresponded to major changes in tissue autofluorescence spectra caused by ablation. Specifically, interested readers are referred to Ref. 10, Fig. 4(c), and Ref. 11, Fig. 4(a) for ablation-induced spectral changes in porcine and human atrial tissue, respectively. The first range of 420 to 510 nm corresponds to decreased signal intensity associated with ablation-induced loss of NADH autofluorescence. The second range of 610 to 630 nm corresponds to increased signal intensity caused by ablation-induced tissue scattering. Ranges 3 and 4 have much less amount of returning light and therefore are less significant. This biological explanation fits well with our observation that SN in green-shaded area of Fig. 13, in which the divider 1 ranges from 510 to 530 nm, includes most high-accuracy combinations. The 510 to 530 nm is, in fact, where ablation-induced change in autofluorescence changes sign from negative to positive.

5 Conclusions

As the most common sustained arrhythmia, AF is expected to affect more than 10 million people by 2050.[1] Our group is developing imaging tools for real-time visualization of ablated tissue. The long-term goal of our studies is to help develop an intracardiac aHSI catheter (Fig. 17) that can improve the success rate of RF treatment, reduce the incidence of AF recurrence, and help to avoid retreatment of the previously ablated tissue.

We have recently demonstrated the ability of autofluorescence hyperspectral imaging to reveal ablated tissue using linear unmixing protocols.[10–14] Here, we have shown that k-means, an approach that does not require *a priori* knowledge of tissue spectra, can be also an effective means to detect lesions from aHSI hypercubes. The average accuracy for detection by k-means ($k = 10$) using 31 features was about 74% when compared to reference images. Second, we have also demonstrated that the number of spectral bands (which are referred to as "features") can be reduced (by grouping them) without significantly affecting lesion detection accuracy. Specifically, we show that by using the best four grouped features, the accuracy of lesion identification was about 94% of that using 31 features. The time cost of four-feature clustering was about 40% of that of the 31-feature clustering, demonstrating that four-feature grouping can speed up acquisition and processing. From an instrumentation point of view, by using a limited number of features one is able to combine multiple spectral bands into one spectrally wide band. This is extremely beneficial for low-light applications such as implementation of aHSI via catheter access.

6 Appendix

Abbreviations/symbols used in this paper are shown in Table 4 and the methods of this study and their interconnections (expanded version of Fig. 1): red lines denote the lesion-detection system in use; green lines define how the best k is found for k-means; blue lines define how the best four-feature groups are

Table 4 Abbreviations/symbols used in this paper and in Fig. 18.

Abbreviation/symbol	Full name/explanation
TR	TTC-reference: the image of stained tissue with TTC, which provides reliable identification of lesions and their boundaries.
GUR	Gray-unmixing-reference: the lesion component image obtained using linear unmixing of a 31-band hypercube as a reference image.
BUR; I_{Ref}	Bi(nary)-unmixing-reference: convert GUR from gray-level image to binary image by Otsu's method.
B31R; $I_{31\text{Rlt}}$	Bi(nary)-31(feature)-result: binary image of lesion detection obtained by k-means clustering using 31 features.
B4R; $I_{4\text{Rlt}}$	Bi(nary)-4(feature)-result: binary image of lesion detection obtained by k-means clustering using grouped four features.
SN	Assigned serial number to each combination of four groups.

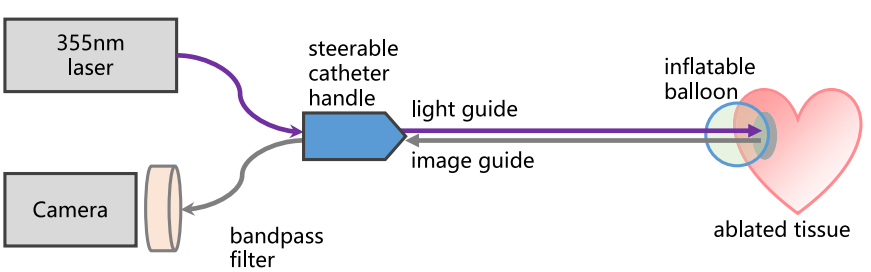

Fig. 17 Proposed concept of acquiring hyperspectral imaging data from the heart.

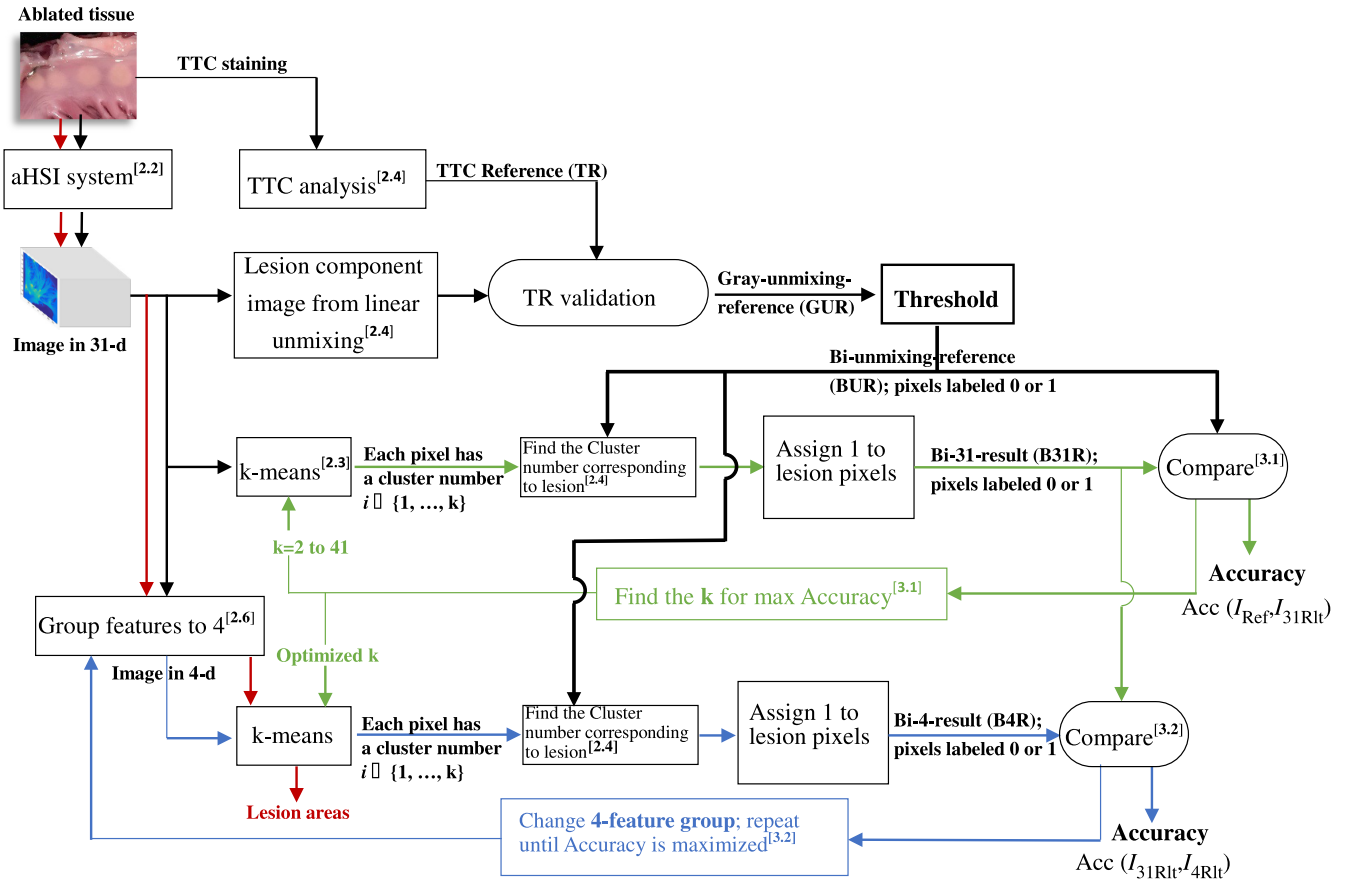

Fig. 18 The methods of this study and their interconnections (expanded version of Fig. 1): red lines denote the lesion-detection system in use; green lines define how the best *k* is found for *k*-means; blue lines define how the best four-feature groups are found; and black lines denote the reference data. Superscripts indicate the sections of the text in which the process is described.

found; and black lines denote the reference data. Superscripts indicate the sections of the text in which the process is described.

Disclosures

No conflicts of interest, financial or otherwise, are declared by the authors.

Acknowledgments

Support from the National Institutes of Health (R42HL120511) is gratefully acknowledged. We thank Dr. Narine Muselimyan for her help in acquiring data in our experiments.

References

1. S. Deshpande, J. Catanzaro, and S. Wann, "Atrial fibrillation: prevalence and scope of the problem," *Card. Electrophysiol. Clin.* **6**, 1–4 (2014).
2. S. Chen et al., "Initiation of atrial fibrillation by ectopic beats originating from the pulmonary veins: electrophysiological characteristics, pharmacological responses, and effects of radiofrequency ablation," *Circulation* **100**, 1879–1886 (1999).
3. R. Cappato et al., "Updated worldwide survey on the methods, efficacy, and safety of catheter ablation for human atrial fibrillation," *Circ. Arrhythm. Electrophysiol.* **3**, 32–38 (2010).
4. A. Arujuna et al., "Acute pulmonary vein isolation is achieved by a combination of reversible and irreversible atrial injury after catheter ablation: evidence from magnetic resonance imaging," *Circ. Arrhythm. Electrophysiol.* **5**, 691–700 (2012).
5. F. Ouyang et al., "Long-term results of catheter ablation in paroxysmal atrial fibrillation: lessons from a 5-year follow-up," *Circulation* **122**, 2368–2377 (2010).
6. H. Oral et al., "Clinical significance of early recurrences of atrial fibrillation after pulmonary vein isolation," *J. Am. Coll. Cardiol.* **40**, 100–104 (2002).
7. A. Wokhlu et al., "Long-term outcome of atrial fibrillation ablation: impact and predictors of very late recurrence," *J. Cardiovasc. Electrophysiol.* **21**, 1071–1078 (2010).
8. R. Ranjan et al., "Identification and acute targeting of gaps in atrial ablation lesion sets using a real-time magnetic resonance imaging system," *Circ. Arrhythm. Electrophysiol.* **5**, 1130–1135 (2012).
9. P. M. Kistler et al., "The impact of CT image integration into an electroanatomic mapping system on clinical outcomes of catheter ablation of atrial fibrillation," *J. Cardiovasc. Electrophysiol.* **17**, 1093–1101 (2006).
10. D. A. Gil et al., "Autofluorescence hyperspectral imaging of radiofrequency ablation lesions in porcine cardiac tissue," *J. Biophotonics* **10**, 1008–1017 (2017).
11. N. Muselimyan et al., "Seeing the invisible: revealing atrial ablation lesions using hyperspectral imaging approach," *PLoS One* **11**, e0167760 (2016).
12. H. Asfour et al., "Optimization of wavelength selection for multispectral image acquisition: a case study of atrial ablation lesions," *Biomed. Opt. Express* **9**, 2189–2204 (2018).
13. L. M. Swift et al., "Hyperspectral imaging for label-free in vivo identification of myocardial scars and sites of radiofrequency ablation lesions," *Hear. Rhythm* **15**(4), 564–575 (2018).

14. N. Muselimyan et al., "Anatomical and optical properties of atrial tissue: search for a suitable animal model," *Cardiovasc. Eng. Technol.* **8**, 505–514 (2017).
15. M. Mercader et al., "Optical tissue interrogation catheter that provides real-time monitoring of catheter-tissue contact and RF lesion progression using NADH fluorescence," *EP Eur.* **18**, i27 (2016).
16. J. Koruth et al., "Direct assessment of catheter-tissue contact and RF lesion formation: a novel approach using endogenous NADH fluorescence," *Hear. Rhythm* **12**(5), S111 (2015).
17. "User's manual for Nuance 3.0.2," p. 104, Perkin-Elmer, Inc., Hopkinton, Massachusetts (2012).
18. T. Hastie, R. Tibshirani, and J. Friedman, Chapter 14 in *The Elements of Statistical Learning*, 2nd ed., Springer, New York (2009).
19. G. James et al., *An Introduction to Statistical Learning*, Vol. **112**, Springer, New York (2013).
20. K. Yamanishi et al., "On-line unsupervised outlier detection using finite mixtures with discounting learning algorithms," *Data Min. Knowl. Discovery* **8**, 275–300 (2004).
21. C. Guo, S. Mita, and D. McAllester, "Robust road detection and tracking in challenging scenarios based on Markov random fields with unsupervised learning," *IEEE Trans. Intell. Transp. Syst.* **13**, 1338–1354 (2012).
22. M. Ranzato et al., "Unsupervised learning of invariant feature hierarchies with applications to object recognition," in *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 1–8 (2007).
23. R. Zhao, W. Ouyang, and X. Wang, "Unsupervised salience learning for person re-identification," in *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 3586–3593 (2013).
24. P. Perconti and M. H. Loew, "Salience measure for assessing scale-based features in mammograms," *J. Opt. Soc. Am. A* **24**, B81–B90 (2007).
25. X. Hou and L. Zhang, "Saliency detection: a spectral residual approach," in *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 1–8 (2007).
26. J. Kim et al., "Salient region detection via high-dimensional color transform," in *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 883–890 (2014).
27. J. A. Hartigan and M. A. Wong, "Algorithm AS 136: a k-means clustering algorithm," *J. R. Stat. Soc. Ser. C* **28**, 100–108 (1979).
28. H.-T. Lim and V. M. Murukeshan, "A four-dimensional snapshot hyperspectral video-endoscope for bio-imaging applications," *Sci. Rep.* **6**, 24044 (2016).
29. S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory* **28**, 129–137 (1982).
30. N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst. Man. Cybern.* **9**, 62–66 (1979).
31. H. Xu and B. W. Rice, "In-vivo fluorescence imaging with a multivariate curve resolution spectral unmixing technique," *J. Biomed. Opt.* **14**, 064011 (2009).

Shuyue Guan is a PhD candidate in biomedical engineering at George Washington University. His primary research interests are the applications of machine learning technologies to solve problems concerning image analysis. His current studies are the ablated tissues (lesion) detection via hyperspectral imaging, and deep-learning based medical imaging. He has published 11 papers in the field of image processing and medical image analysis and exhibited his works in 10 exhibitions during his doctoral program.

Huda Asfour gained her BSc and PhD degrees in electrical engineering in 2006 and 2012 from the George Washington University (GW). She completed her postdoctoral training at Duke University before returning to GW. Currently, she serves as a research scientist and adjunct professor there. A multidisciplinary researcher with training in optical imaging, tissue engineering and signal and image processing, her main research interests are developing software and hardware solutions for advancing cardiac imaging.

Narine Sarvazyan is a professor of Physiology at George Washington University School of Medicine. She received degrees in physics and biophysics from the Physics Department of Moscow State University, and her PhD in biological sciences from the Institute of Experimental Biology, Yerevan, Armenia (1991). Her research interests encompass a variety of cardiovascular topics including cardiotoxicity of phthalates and anthracyclines, origins of ectopic arrhythmias, stem-cell based cardiac regeneration and new imaging modalities to examine pathophysiological tissue states.

Murray Loew is a professor in the Department of Biomedical Engineering at George Washington University, Washington, DC, and director of the Medical Imaging and Image Analysis Laboratory. His interests include the development and application of image classification, thermal and hyperspectral imaging, and image fusion techniques, using machine learning for disease detection and outcome prediction. He is a Fellow of SPIE, IEEE, and AIMBE, and the inaugural Fulbright U.S.-Australia Distinguished Chair, Advanced Science and Technology (2014).