PAPER • OPEN ACCESS

Speeded-Up Robust Features-based image mosaic method for large-scale microscopic hyperspectral pathological imaging

View the article online for updates and enhancements.

Speeded-Up Robust Features-based image mosaic method for large-scale microscopic hyperspectral pathological imaging

Qing Zhang[1], Li Sun[1], Jiangang Chen[1], Mei Zhou[1,2], Menghan Hu[1], Ying Wen[1] and Qingli Li[1,2,3] (iD)

[1] Shanghai Key Laboratory of Multidimensional Information Processing, East China Normal University, Shanghai, People's Republic of China
[2] Engineering Center of SHMEC for Space Information and GNSS, Shanghai, People's Republic of China
[3] Engineering Research Center of Nanophotonics & Advanced Instrument, Ministry of Education, East China Normal University, Shanghai, People's Republic of China

E-mail: qlli@cs.ecnu.edu.cn

Abstract

Microscopic hyperspectral imaging technology has been widely used in pathological analysis as it can obtain both spatial and spectral information of samples. However, most hyperspectral imaging systems can only capture images in a single field of view. Therefore, an image mosaic is one of the most important steps in a large-scale microscopic hyperspectral imaging system. This paper proposes a microscopic hyperspectral image (HSI) mosaic method based on Speeded-Up Robust Features (SURF) and linear synthesis to achieve large-scale HSIs. In contrast to other SURF-based image mosaic methods, the proposed method leverages both image content and coordinate information to improve the accuracy and stability of the image mosaic. In addition, multiple bands of HSIs with different texture information and grayscale are applied in image matching to take full advantage of spatial redundancy. Simultaneously, a blank microscopic HSI screening method is introduced in this paper to pick out a clearer blank image for better preprocessing, i.e. removing interference in the optical path and the interference of dust on slides. Finally, the preprocessed images are synthesized by linear-based synthesis methods due to their simple synthesis structure and better universality. Additionally, a file format, i.e. hyperslide, is defined for large-scale HSIs and can be browsed with HyperViewer software. Experimental results show that the proposed microscopic HSI mosaic method can obtain high-quality large-scale microscopic HSIs of tissue sections.

Keywords: microscopic hyperspectral imaging, large-scale imaging, image mosaic, pathology

(Some figures may appear in colour only in the online journal)

1. Introduction

Pathological diagnosis is of great importance in tumor identification, and microscopic imaging techniques have been widely used for this purpose to provide the specific structures of biological cells, organelles and tissues [1–3]. With the development of artificial intelligence as well as novel cheaper and more powerful technologies, increasingly large and complex neural networks are being presented online and new requirements of turning single-field imaging into whole-slide imaging are emerging for pathological images [4]. Individual images, i.e. tiles, usually represent very small areas because they are scanned at very high microscope magnifications for clear texture information, which will directly affect the observation and analysis of sections [5]. In contrast, if the section is scanned at low magnification to obtain an overall image, it is easier to observe the distribution pattern. Nonetheless, the low resolution of these images seriously restricts feature extraction. Therefore, a natural solution to this problem is to leverage computational tools to assemble several individual tiles to form a high-resolution large-scale image [6]. Subsequently, several large-scale pathological databases have also been introduced for the training and testing of neural networks [7, 8]. In this case, a whole-slide image (WSI) with high resolution (submicrometer) will assist in clinical diagnosis by giving access to both the detailed structure and overall preview of tissue slide images within a relatively short time [9, 10]. Automated whole-slide imaging scanners have been generally used in education, medicine and institutes [11]. For example, Kawano *et al* proposed a microscopic whole-slide imaging system based on color darkfield internal reflection illumination for biological applications [12], and Guo *et al* reported a real-time autofocusing-strategy-based whole-slide imaging system with low cost and high throughput [13].

Autofocus and image mosaics are both important steps in whole-slide imaging systems or large-scale pathological imaging systems. To date, several researchers have proposed various image mosaic algorithms mainly based on Speeded-Up Robust Features (SURF) and scale-invariant feature transform (SIFT) to acquire large-scale pathological images [14]. Shan *et al* leveraged SURF in a mosaic for two adjacent images and then matched eigen points via the Bray–Curtis similarity [15]. Pang *et al* solved the image mosaic problem of blood cells using a SIFT–Gauss splicing fusion method [16]. However, current pathological diagnosis is mostly based on color images and their limited information obstructs improvements in diagnostic accuracy. Therefore, the microscopic hyperspectral image (HSI) technique is incorporated into pathological diagnosis such as red blood cell count automation, lymphoblastic leukemia cell identification, chemical selective analysis of malaria-infected red blood cells, etc [17–19], whereas studies on HSI-technology-based large-scale pathological imaging are still rare. Consequently, after the study of the autofocus algorithm, it is necessary to further study the corresponding mosaic algorithm to obtain high-quality microscopic hyperspectral pathological image data for pathological diagnosis.

The reconstruction of microscopic hyperspectral WSIs is of great importance but difficult because the mosaic algorithm has to simultaneously preserve both spatial information and spectral information. Therefore, some researchers have implemented specific precise devices to generate the WSIs, such as prism-based slit-array dispersion and digital micromirror device-based spatial light modulators [20, 21]. Additionally, some microscopic HSI mosaic algorithms, usually consisting of image preprocessing, image registration and image fusion, have also been put forward to obtain microscopic hyperspectral WSIs. The preprocessing of microscopic HSIs is usually implemented by dividing a blank image captured on the same slide with the same light source. This will remove interference from the dirty spots on the slide and optical path. Image registration is a fundamental component in medical image processing and it can be formulated as an optimization problem, finding the optimal spatial transformation model between two or more images [22]. Image registration is generally composed of a coordinate transformation model, optimization model, multiresolution strategy, image interpolation method and cost function [23]. The methods of image registration are typically classified into two groups: nonfeature-based and feature-based [24]. Nonfeature-based methods are sensitive to image transformation and are computationally complex tasks [24], whereas feature-based methods are faster and more robust in relation to noise and large deformations. The frequently utilized features contain points (region corners, points with high curvature), lines (section boundaries, blood vessels, fibrous tissue) and significant regions (tumor regions, hyperplasia regions, cells) [25]. There are several feature-based image registration methods proposed, such as SIFT and SURF [26]. SIFT leverages a high-dimensional vector to describe a feature point and each feature point is independent of image size and rotation with good tolerance to light and noise [27, 28]. Al-khafaji proposed a spectral–spatial SIFT method to extract spectral–spatial scale-invariant features for HSI matching [29]. SURF is similar to SIFT with respect to feature point matching, but it can extract many robust local features more rapidly thanks to two-dimensional Haar wavelet responses, integral images and scale space technique [30–32].

Image fusion is also a fundamental image processing technique that stitches several images to obtain a large-scale image for better identification, detection and segmentation of natural and artificial objects [33, 34]. Large-scale microscopic high-spatial-resolution HSIs (HRHSIs) are now mostly obtained by fusing low-spatial-resolution HSIs (LRHSIs) and high-spatial-resolution multispectral images (HRMSIs) [35–37]. These methods are generally classified into three groups: extensions of sharpening, model-based and low-rank structure-based fusion methods [38]. The span-sharpening methods typically utilize the spatial information of a panchromatic image to sharpen the multispectral images (MSIs) to generate HRHSIs [39]. However, some undesirable effects may occur because of spectral range differences between LRHSIs and HRMSIs [40]. The second method, model-based, is the most popular method in HSI fusion problems. In this method, LRHSIs and HRMSIs are respectively regarded as the decomposition part of HRHSIs in the spatial and spectral domains through degradation matrices [41]. Nevertheless, these large matrices in both the spectral and spatial domains will lead to

higher space complexity. Regarding the third fusion method, the low-rank properties are preserved when analyzing the HSIs and then each pixel is processed independently without respect to the inherent spatial relationship [38, 42]. Therefore, this method results in distortions of the spectral structure and spatial details as well as the spectral features. The above three fusion methods are widely applied in HSI and MSI fusion and they have obtained good results in generating HRHSIs. However, these methods require both HSIs and MSIs, which is too time-consuming and expensive in memory during image acquisition.

In this paper, we propose a microscopic HSI mosaic algorithm based on SURF and linear synthesis to obtain large-scale microscopic HSIs. These images were captured by a homemade acousto-optic tunable filter (AOTF) hyperspectral imaging system with an autofocus algorithm considering wavelength changes [43, 44]. In this algorithm, the content-based mosaic method is repeatedly leveraged to captured microscopic HSIs rather than both HSIs and MSIs. Moreover, this method is supplemented with the position-based mosaic method so that the mosaic HSIs are generated with higher accuracy and stability. In addition, there are differences in texture and grayscale between different wavelengths and the spatial redundancy-based multi-wavelength characteristics are applied to the microscopic HSI mosaic. Experimental results show that the proposed algorithm can obtain high-quality large-scale microscopic HSIs, which is useful for building HSI pathological databases.

2. Methods

Microscopic image mosaic algorithms have been widely used in large-scale color image mosaics, and most of these algorithms can be divided into two categories: multi-image mosaic with overlapping regions and multi-image mosaic without overlapping regions. As shown in figure 1, the application of a multi-image mosaic with overlapping regions requires several choices to be made, such as multiple images acquired in a certain region, the overlapping area between each pair of images and a suitable size of overlapping area.

Given two horizontally adjacent images with overlapping regions, the ratio of the width of the overlapping regions to the width of the $w \times h$ image is assumed to be α. The similarity of eigenvectors is calculated to obtain the shift of each image from the combined image, i.e. displacement vector $(\Delta w, \Delta h)$. In ideal circumstances, if the left image stays still, the displacement vector of the left image is (0,0) and that of the right image is $(w \times \alpha, 0)$. Similarly, when it comes to two vertically adjacent images, if the above image remains still, the displacement vector of the above image is (0,0) and that of the below image is $(0, h \times \alpha)$. As a result, two images are composited together in figure 2 based on the displacement vectors to generate a large image with size $(w + \Delta w_1 + \Delta w_2, h + \Delta h_1 + \Delta h_2)$.

For multiple images without overlapping regions, extra information is needed for the image mosaic. It is usually implemented by recording displacement information, e.g. absolute
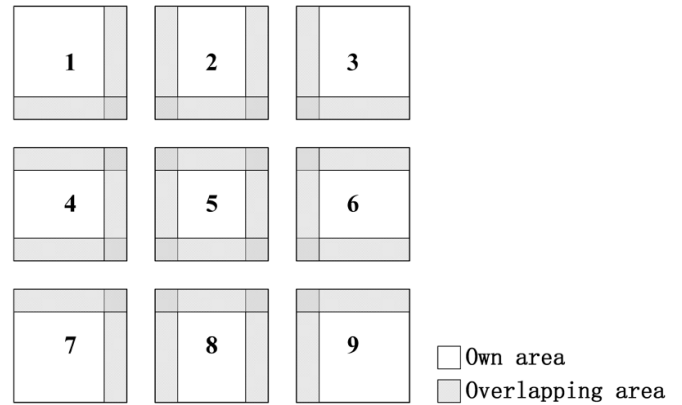

Figure 1. Adjacent images are captured with overlapping regions. Shadowed areas represent overlapping regions and blank areas represent non-overlapping regions.
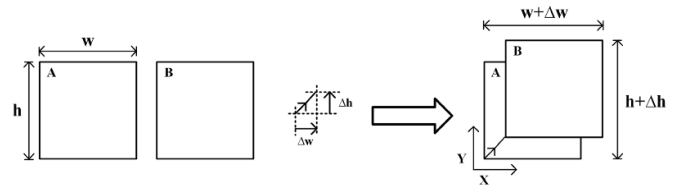

Figure 2. Two adjacent images are composited based on the displacement vectors. Δw and Δh are displacement vectors in the horizontal and vertical directions respectively.

position of the triaxial electric loading platform when acquiring the current image or relative shift to the previous position of the triaxial electric loading platform. The displacement of the triaxial electric loading platform is precisely controlled to make sure that the content of two adjacent images is continuous and there is no gap or a very small gap between the two images. Then, all images are stitched together to obtain the large-scale image according to the position information. This method has a high requirement of accuracy of the electric platform in the acquisition system, otherwise there will be a large gap in the resultant image.

The above two methods are both adopted in this system, where the multi-image mosaic with overlapping regions dominates and the multi-image mosaic without overlapping regions assists in inspecting and correcting to increase the accuracy of the algorithm. In addition, microscopic HSI is abundant in information, especially in spatial redundancy. Therefore, the image-content-based mosaic method is repeatedly used supplemented by the image-position-based mosaic method to obtain higher accuracy and stability.

2.1. Image registration algorithm

An intersection must exist between SURF eigenvectors of two adjacent images with overlapping regions. The coordinate mapping relation is established for the eigen points with the same eigenvector in the two images, i.e. offset vector. The images with overlapping regions in this paper are captured under the same light source settings, so some
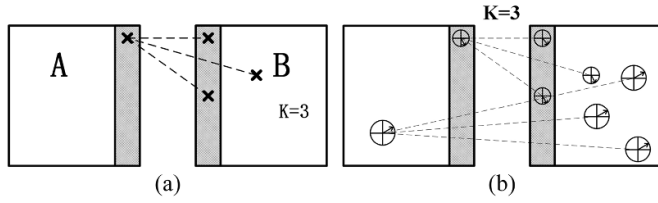
Figure 3. Multi-point matching in KNN feature points.

Figure 4. SURF eigen points: (a) sketch map; (b) actual map.

SURF eigen points have different coordinate information but similar eigenvectors. Simultaneously, the characteristic also lies in SURF eigen points of overlapping regions and non-overlapping regions. As shown in figure 3(a), the *k*-nearest-neighbor (KNN) [45, 46] combined with random sample consensus (RANSAC) [47, 48] is leveraged to generate and screen eigen points to receive a pair of effective eigen points. Compared with optimum matching methods, KNN calculates the first *K* eigen points and then discards some of them based on the situation. Optimum matching methods are much easier, whereas a correct eigen point may be missed because there are many similar tissues in microscopic physical images.

To generate correct eigen points, SURF eigen points of both image A and image B are calculated through the SURF algorithm. *N* SURF eigen points in image A are recorded as $\{\vec{v}_{A,1}, \vec{v}_{A,2}, \cdots, \vec{v}_{A,i}, \cdots, \vec{v}_{A,N}\}, 1 \leqslant i \leqslant N$ and in image B as $\{\vec{v}_{B,1}, \vec{v}_{B,2}, \cdots, \vec{v}_{B,j}, \cdots, \vec{v}_{B,M}\}, 1 \leqslant j \leqslant M$. Every $\vec{v}$ is a 64-dimensional vector, i.e. $\vec{v}_{A,i} = \{A_1, A_2, \cdots, A_{64}\}$ and $\vec{v}_{B,j} = \{B_1, B_2, \cdots, B_{64}\}$. The similarity of all eigen points in image A and in image B, containing $N \times M$ eigen point pairs, is calculated in (1)

$$\begin{aligned} &\left\{ \left(\vec{v}_{A,1}, \vec{v}_{B,N_{1,1}}\right), \left(\vec{v}_{A,1}, \vec{v}_{B,N_{1,2}}\right), \cdots, \left(\vec{v}_{A,1}, \vec{v}_{B,N_{1,j}}\right), \cdots, \right. \\ &\left(\vec{v}_{A,1}, \vec{v}_{B,N_{1,M}}\right), \left(\vec{v}_{A,i}, \vec{v}_{B,N_{i,1}}\right), \left(\vec{v}_{A,i}, \vec{v}_{B,N_{i,2}}\right), \cdots, \left(\vec{v}_{A,i}, \vec{v}_{B,N_{i,j}}\right), \\ &\cdots, \left(\vec{v}_{A,i}, \vec{v}_{B,N_{i,M}}\right), \left(\vec{v}_{A,N}, \vec{v}_{B,N_{N,1}}\right), \left(\vec{v}_{A,N}, \vec{v}_{B,N_{N,2}}\right), \cdots, \\ &\left. \left(\vec{v}_{A,N}, \vec{v}_{B,N_{N,j}}\right), \cdots, \left(\vec{v}_{A,N}, \vec{v}_{B,N_{N,M}}\right) \right\}. \end{aligned}$$
(1)

It is shown in figure 3(b) that multiple point matching may appear in KNN. Considering that the image shifts linearly, the correct connection is A and D, whereas B and C should be removed. Therefore, RANSAC is utilized to remove unconscionable eigen points by means of iteration to help with the generation of the correct offset vector during the scan of eigen point pairs. The eigenvectors of SURF eigen points are similar to each other, so a set of eigen points $\{(\vec{v}_{A,i}, \vec{v}_{B,N_{i,1}}), (\vec{v}_{A,i}, \vec{v}_{B,N_{i,2}}), \ldots, (\vec{v}_{A,i}, \vec{v}_{B,N_{i,j}}), \ldots, (\vec{v}_{A,i}, \vec{v}_{B,N_{i,M}})\}$ is generated in this system. SURF eigen point $\vec{v}$ contains the sub-pixel coordinate information, i.e. $(x_{\vec{v}}, y_{\vec{v}})$, which is important information in extracting SURF eigen points. For one pair of SURF eigen points $(\vec{v}_{A,i}, \vec{v}_{B,N_{i,j}})$, the coordinate information of $\vec{v}_{A,i}$ is $(x_{\vec{v}_{A,i}}, y_{\vec{v}_{A,i}})$ and that of $\vec{v}_{B,N_{i,j}}$ is $(x_{\vec{v}_{B,N_{i,j}}}, y_{\vec{v}_{B,N_{i,j}}})$, as shown in figure 4.

One pair of SURF eigen points can be formulated as a two-dimensional vector $\vec{p}_{i,j}$ in (2). Most of the two-dimensional vectors point in the same direction and few of them diverge, and thus they can be eliminated by RANSAC
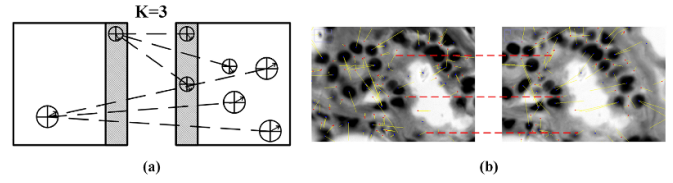
$$\vec{p}_{i,j} = \left(x_{\vec{v}_{A,i}} - x_{\vec{v}_{B,N_{i,j}}}, y_{\vec{v}_{A,i}} - y_{\vec{v}_{B,N_{i,j}}}\right).$$
(2)

It is assumed in this system that a linear translation model finally fits a linear translation vector, i.e. the offset vector of the two images $\vec{p}_{\text{final}}$. Given a threshold $\delta \in (0, 1)$ and an iteration *k*, all the $N \times M$ SURF eigen points will satisfy this model only if there are at least $\delta \times N \times M$ SURF eigen points satisfying it. The specific process of RANSAC screening is as follows:

(1) Choose *m* eigen point pairs from $N \times M$ SURF eigen point pairs to obtain the initialized linear translation vector $\vec{p}_{\text{init}}$.
(2) Compute the remaining $(N \times M - m)$ SURF eigen point pairs. If the two-dimensional vector $\vec{p}_{i,j}$ of one pair of eigen points $(\vec{v}_{A,i}, \vec{v}_{B,N_{i,j}})$ satisfies (3), this pair of eigen points is a reasonable value and vice versa

$$|\vec{p}_{i,j} - \vec{p}_{\text{init}}| \leqslant \varepsilon$$
(3)

where ε is the internal threshold in the linear translation model, i.e. the degree to which the two-dimensional vector of each eigen point pair is shifted from the initial linear translation vector.
(3) If the number of all reasonable SURF eigen point pairs is larger than $\delta \times N \times M$, the final linear translation vector $\vec{p}_{\text{final}}$ is calculated based on these eigen point pairs.
(4) If the number of all reasonable SURF eigen point pairs is smaller than or equal to $\delta \times N \times M$, go back to step 1 and the iteration number plus one. If the maximum iteration number is exceeded, $\vec{p}_{\text{final}}$ is generated from the model of the first *k* models that fits the data best.

In summary, RANSAC scans all the eigen point pairs and removes unusual SURF eigen point pairs. This method will preserve large numbers of reasonable eigen point pairs to obtain the relative displacement of two images $\vec{p}_{\text{final}}$.

2.2. Image fusion algorithm

Microscopic HSI includes more abundant spatial information compared with signal gray imaging, which is taken into consideration to receive microscopic HSI fusion with higher accuracy. Several images of the same field of view $\{I_1, I_2, \cdots, I_N\}$ are acquired at different wavelengths $\{\lambda_1, \lambda_2, \cdots, \lambda_N\}$ (*N* is the number of wavelengths in microscopic HSI) and there are texture feature and gray differences due to the different wavelengths, as shown in figure 5.

Considering that the differences between these images lead to relatively large differences between SURF eigen points, a
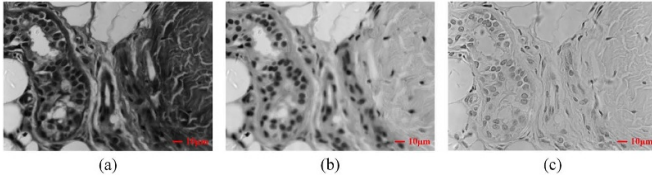
Figure 5. Gray images of the same field of view at different wavelengths in melanoma samples (40×): (a) 510 nm; (b) 606 nm; (c) 703 nm.

multi-wavelength image matching method based on the spatial redundancy feature is adopted in this system and then the images are fused based on the location information to obtain the offset vectors of two adjacent images.

The microscopic HSI data consists of N gray images $\{I_1, I_2, \cdots, I_N\}$ at different wavelengths $\{\lambda_1, \lambda_2, \cdots, \lambda_N\}$ and its physical position in the imaging system is represented by $\{x_H, y_H, z_H\}$. The calculation steps are as follows:

(a) Calculate the initial offset vector $\vec{p}_{\text{init}}$.

Extract grayscale images $I_{1,n}, I_{2,n}$ at the same wavelength $\lambda_{\tilde{n}}$ from two adjacent microscopic HSIs H_1, H_2 and then calculate the initial offset vector $\vec{p}_{\text{init}}$ as demonstrated in section 2.1. The wavelength $\lambda_{\tilde{n}}$ has to satisfy (4)

$$\lambda_{\tilde{n}} = (\lambda_1 + \lambda_2)/2. \tag{4}$$

If $\lambda_{\tilde{n}} \notin \{\lambda_1, \lambda_2, \cdots, \lambda_N\}$, $\lambda_{\tilde{n}}$ must be approximated. $\forall n \in (1, N), \{\lambda_1, \lambda_2, \cdots, \lambda_N\}$, if $|\lambda_n - \lambda_{\tilde{n}}|$ can obtain the minimum, λ_n is the approximate value of $\lambda_{\tilde{n}}$.

(b) Calculate the offset correction vector $\vec{p}_{\text{pos}}$.

The physical coordinate information of each microscopic HSI is recorded in the microscopic HSI system. The physical coordinates of two adjacent microscopic HSIs H_1, H_2 are recorded as $\{x_1, y_1, z_1\}$ and $\{x_2, y_2, z_2\}$. The coordinate unit is millimeters with four decimal places reserved and the highest precision is $0.1\,\mu\text{m}$. First, calculate the pixel numbers under different objectives based on the magnification factor per unit size α. There are five magnification factors $\{\alpha_4, \alpha_{10}, \alpha_{20}, \alpha_{40}, \alpha_{100}\}$ which correspond to $4\times, 10\times, 20\times, 40\times, 100\times$ objectives. Then, calculate the offset in physical position (dx, dy, dz) of the two images as formulated in (5)

$$\begin{cases} dx = x_1 - x_2 \\ dy = y_1 - y_2 \\ dz = z_1 - z_2 \end{cases}. \tag{5}$$

The position of the image on the Z axis is not required in the image mosaic, so this paper only calculates the pixel offset on the X-axis and Y-axis, i.e. offset correction vector $\vec{p}_{\text{pos}} = (p_x, p_y)$

$$p_x = dx \times \alpha, p_y = dy \times \alpha. \tag{6}$$

(c) Calculate the fusion offset vector $\vec{p}_{\text{fusion}}$.

Any two grayscale images $\{I_{1,n}, I_{2,n}, 1 \leqslant n \leqslant N\}$ of two adjacent HSIs at the same wavelength and another two grayscale images $\{I_{1,m}, I_{2,m}, 1 \leqslant m \leqslant N, m \neq n\}$ at another wavelength have the same offset vector. However, in a practical imaging system, the images will have different grayscale distributions, contrast ratios and texture features because of the different incident wavelengths, resulting in deviations of offset vectors. Hence, multi-wavelength grayscale images are adopted to calculate the fusion offset vector through linear fusion on account of wavelength information.

The spectral range of microscopic HSI is certain, i.e. $(\lambda_{\text{start}}, \lambda_{\text{end}})$ and the interval between wavelengths λ_{step} is formulated in (7) if the image contains N bands. When M bands of the image are extracted, the interval between wavelengths turns into as calculated in (8)

$$\lambda_{\text{step}} = (\lambda_{\text{end}} - \lambda_{\text{start}})/(N-1) \tag{7}$$

$$\lambda_{\_\text{step}} = (\lambda_{\text{end}} - \lambda_{\text{start}})/(M-1). \tag{8}$$

Ideally, the microscopic HSI of M bands is suitable only when and satisfy (9)

$$\lambda_{\_\text{step}} = K\lambda_{\text{step}}, K \in Z^+, M < N. \tag{9}$$

M suitable wavelengths $\{\lambda_{l_1}, \lambda_{l_2}, \cdots, \lambda_{l_M}\}$ are as follows:

$$\lambda_{l_i} = \lambda_{\text{start}} + i\lambda_{\_\text{step}}, 1 \leqslant i \leqslant M. \tag{10}$$

If λ_{step} and $\lambda_{\_\text{step}}$ satisfy (9), M suitable wavelengths will satisfy (11). Otherwise, find the proximate wavelengths from N wavelengths. $\forall \lambda_{l_i} \in \{\lambda_{l_1}, \lambda_{l_2}, \cdots, \lambda_{l_M}\}, \forall \lambda_n \in \{\lambda_1, \lambda_2, \cdots, \lambda_N\}$, if $|\lambda_{l_i} - \lambda_n|$ gets the minimum, λ_n will be the approximate value of λ_{l_i}

$$\{\lambda_{l_1}, \lambda_{l_2}, \cdots, \lambda_{l_M}\} \subseteq \{\lambda_1, \lambda_2, \cdots, \lambda_N\}. \tag{11}$$

M grayscale images of the two adjacent microscopic HSIs $\{I_{1,l_1}, I_{1,l_2}, \cdots, I_{1,l_M}\}, \{I_{2,l_1}, I_{2,l_2}, \cdots, I_{2,l_M}\}$ at corresponding wavelengths $\{\lambda_{l_1}, \lambda_{l_2}, \cdots, \lambda_{l_M}\}$ are extracted from the initial images. Then, a set of offset vectors $\{\vec{p}_{l_1}, \vec{p}_{l_2}, \cdots, \vec{p}_{l_M}\}$ of the image pairs $\{I_{1,l_1}, I_{2,l_1}\}, \{I_{1,l_1}, I_{2,l_2}\}, \cdots, \{I_{1,l_M}, I_{2,l_M}\}$ is calculated on the basis of 2.1. Outliers may exist in the set of offset vectors $\{\vec{p}_{l_1}, \vec{p}_{l_2}, \cdots, \vec{p}_{l_M}\}$ in a practical imaging system, so the position information is exploited to correct the vectors. If offset vector $\vec{p}_{l_i}, 1 \leqslant i \leqslant M$ and offset correction vector $\vec{p}_{\text{pos}}$ satisfy (12), $\vec{p}_{l_i}$ is considered as the reasonable offset vector, otherwise the offset vector will be removed

$$0.9\vec{p}_{\text{pos}} \leqslant \vec{p}_{l_i} \leqslant 1.1\vec{p}_{\text{pos}}. \tag{12}$$

The original offset vectors $\{\vec{p}_{l_1}, \vec{p}_{l_2}, \cdots, \vec{p}_{l_M}\}$ turn into $\{\vec{p}_{l_1}, \vec{p}_{l_2}, \cdots, \vec{p}_{l_Q}\}$ after position correction and these Q offset vectors are then linearly fused to obtain fusion offset vector $\vec{p}_{\text{fusion}}$. The corresponding weights of linear fusion to
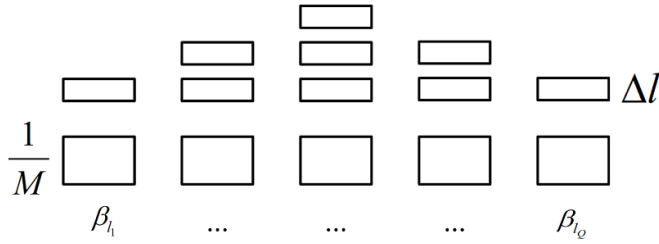
Figure 6. Schematic diagram of the weight of the offset vector.

$\{\vec{p}_{l_1}, \vec{p}_{l_2}, \cdots, \vec{p}_{l_Q}\}$ is $\{\beta_{l_1}, \beta_{l_2}, \cdots, \beta_{l_Q}\}$, representing the contribution of the offset vectors of different wavelength images to the fusion offset vector. In this paper, a linear weight calculation method based on wavelength differentiation is adopted. A certain wavelength is taken as the intermediate point, and the difference between the wavelength of each offset vector and the wavelength of the intermediate point is taken as the weight. Subsequently, the weight of Q offset vectors is obtained by normalization. If $M = Q$, i.e. M offset vectors are all reasonable, all weights will be evenly divided in (13). Otherwise, some offset vectors are removed and the remaining weights are evenly divided as shown in figure 6

$$\beta_{l_1} = \beta_{l_2} = \cdots = \beta_{l_Q} = \frac{1}{Q}, Q = M. \qquad (13)$$

The wavelength interval between offset vectors Δl is defined by the equation

$$E = \begin{cases} 2\left(1 + 2 + \cdots + \frac{Q}{2}\right), Q \text{ is even} \\ 2\left(1 + 2 + \cdots + \frac{Q+1}{2}\right) - \frac{Q+1}{2}, Q \text{ is odd} \end{cases}$$

$$\Delta l = \left(1 - \frac{Q}{M}\right)/E. \qquad (14)$$

The weight value is given as

$$\beta_{l_q} = \begin{cases} \frac{1}{M} + \Delta l \times q, 1 \leqslant q \leqslant \frac{Q}{2} \\ \beta_{l_{\frac{q}{2}}} + \Delta l \times \left(\frac{Q}{2} - q + 1\right), \frac{Q}{2} \leqslant q \leqslant Q \end{cases}. \qquad (15)$$

Finally, the fusion offset vector $\vec{p}_{fusion}$ is calculated as

$$\vec{p}_{fusion} = \sum_{i=1}^{Q} \beta_{l_i} \vec{p}_{l_i}. \qquad (16)$$

(d) Offset the vector of microscopic HSI $\vec{p}_{final}$.

If offset vectors at different wavelengths $\{\vec{p}_{l_1}, \vec{p}_{l_2}, \cdots, \vec{p}_{l_Q}\}$ are with sole error, there will be a gap in the mosaic image. In this circumstance, the fusion offset vector $\vec{p}_{fusion}$ should be corrected again by initial offset vector $\vec{p}_{init}$. Before the correction, the error vector $\vec{p}_{err}$ between the two vectors is calculated as

$$\vec{p}_{err} = \frac{\vec{p}_{fusion} - \vec{p}_{init}}{\vec{p}_{init}}$$

$$\text{err} = |\vec{p}_{err}| = \sqrt{\vec{p}_{err,x}^2 + \vec{p}_{err,y}^2} \qquad (17)$$

where error vector $\vec{p}_{err}$ represents the offset of initial offset vector $\vec{p}_{init}$ from fusion offset vector $\vec{p}_{fusion}$. In view of the higher stability of the multi-wavelength fused offset vector $\vec{p}_{fusion}$, the two vectors are averaged when $\vec{p}_{err}$ is very small. The larger the error vector $\vec{p}_{err}$, the smaller the difference between $\vec{p}_{final}$ and $\vec{p}_{fusion}$. By that analogy, when $\vec{p}_{err}$ is too large, $\vec{p}_{fusion}$ is adopted as $\vec{p}_{final}$ and $\vec{p}_{init}$ is considered to have a large error

$$\vec{p}_{final} = \chi \vec{p}_{final} + (1 - \chi) \vec{p}_{fusion}$$

$$\chi = \begin{cases} \frac{1}{2}, 0 \leqslant \text{err} < 0.1 \\ 0.7 - \text{err} \times 2, 0.1 \leqslant \text{err} < 0.35 \\ 0, \text{err} \geqslant 0.35 \end{cases}. \qquad (18)$$

2.3. Blank microscopic HSI screening algorithm

In the process of large-scale multidimensional data acquisition, a series of microstructure HSIs with large margins will influence the accuracy of the autofocus and image mosaic. Therefore, the proportion of empty area is calculated in this paper. First, two grayscale images $I_{t,n}, I_{b,n}$ of the nearest wavelength λ_n of λ_{ref} are separately extracted from microscopic HSI with histopathological tissue H_t and blank microscopic HSI H_b. The same wavelengths of H_t and H_b are defined as $\{\lambda_1, \lambda_2, \cdots, \lambda_N\}$ and λ_n is defined as

$$\lambda_n = \begin{cases} \lambda_m, |\lambda_m - \lambda_{ref}| = \min_{1 \leqslant i \leqslant N} |\lambda_i - \lambda_{ref}|, \lambda_1 \leqslant \lambda_{ref} \leqslant \lambda_N \\ \lambda_N, \lambda_N < \lambda_{ref} \end{cases}. \qquad (19)$$

Secondly, the ratio image I_r is calculated as

$$I_r = I_{t,n}/I_{b,n}. \qquad (20)$$

The two grayscale images $I_{t,n}, I_{b,n}$ have the same incident light and $I_{b,n}$ has higher mean brightness because there are no histopathological tissues in the blank area. Therefore, the pixel values in the ratio image I_r are all within [0,1] and individual outliers are also set within [0,1].

Finally, the ratio image I_r is transferred to a binary image using its average pixel values and the binary image is transferred into six parts in figure 7 to calculate the ratio of blank area in I_r.

The ratios of blank area in the six parts are respectively calculated as

$$r = \frac{N_{white}}{N_{total}} = \frac{N_{white}}{wh}, w = \frac{W}{3}, h = \frac{H}{2} \qquad (21)$$

where white represents white pixels and total represents total pixels. r is leveraged in the system to optimize autofocus and large-scale image acquisition, reducing the negative impact of large blank areas.

Simultaneously, a blank image is of great importance in microscopic HSI processing because it assists in both spectral correction and removal of intrinsic interference. In addition, if the blank image is disturbed by factors such as dust on the slide, the image quality will deteriorate. Therefore, two blank images H_1, H_2 are taken during image acquisition and then one is screened through a simple evaluation method based on image brightness for image preprocessing.
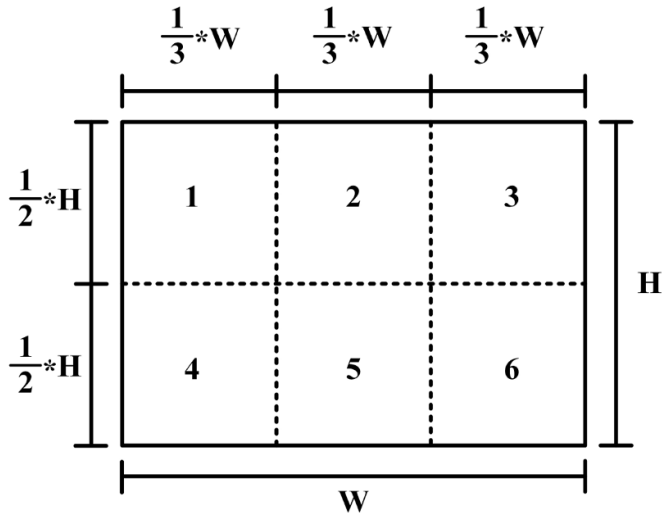
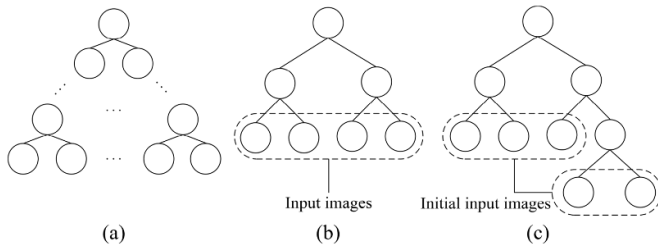**Figure 7.** Area division for the calculation of blank area ratio.

Figure 8. Synthesis tree: (a) synthesis tree of tree-based synthesis strategy; (b) even number of nodes in one layer; (c) odd number of nodes in one layer.

First, two grayscale images $I_{1,n}, I_{2,n}$ at wavelength λ_n, which is calculated in (19), are chosen in the two blank images. Then, the two images are respectively treated as an image with histopathological tissue and a blank image, and the ratios S_1, S_2 of blank area in $I_{1,n}, I_{2,n}$ are calculated in (22) following formula (21). If there is some dust in $I_{1,n}$ and $I_{2,n}$ is clearer, the ratio image S_2 will be determined as blank area. In the end, the clearer blank HSI S_{final} is screened by comparing the total pixel values of S_1 and S_2

$$
\begin{aligned}
S_1 &= \text{Calculate}\left(I_{1,n}/I_{2,n}\right) \\
S_2 &= \text{Calculate}\left(I_{2,n}/I_{1,n}\right)
\end{aligned}
\tag{22}
$$

$$
S_{\text{final}} = \begin{cases} H_2, \text{Sum}\left(S_1\right) \leqslant \text{Sum}\left(S_2\right) \\ H_1, \text{Sum}\left(S_1\right) > \text{Sum}\left(S_2\right) \end{cases}.
\tag{23}
$$

2.4. Microscopic HSI synthesis strategy

Multi-image synthesis can be simplified into multiple two-image synthesis and the process is then iterated in a loop. There are usually two kinds of synthesis strategy: tree-based and linear-based. In a tree-based synthesis strategy, the process of multiple two-image synthesis can be demonstrated as a binary tree, i.e. synthesis tree as shown in figure 8(a). The basic idea of the synthesis tree is to compose a synthesis task with
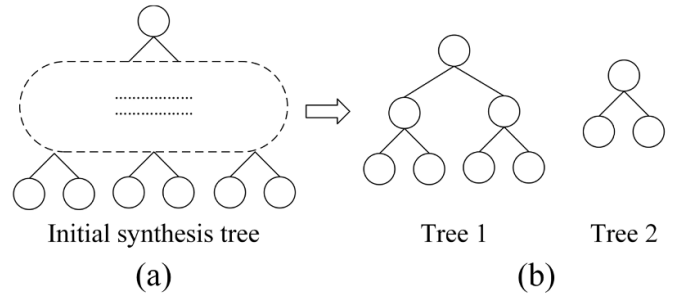

Figure 9. Initial synthesis tree transferred into multiple synthesis trees: (a) synthesis tree of linear-based synthesis strategy; (b) two synthesis trees transformed from the initial synthesis tree.

two child nodes and one parent node, and then gradually form a large-scale image through a bottom-up step-by-step mosaic. The image on the root node is the resultant image, i.e. final large-scale multidimensional image.

In practical experiments, the number of images is random, so two methods are introduced when the number of microscopic HSIs is *N*. One method is to avoid grouping the images, always keep one synthesis tree and adjust it according to the number of nodes in each layer. If the number of nodes *n* in one layer is even as in figure 8(b), the nodes in this layer are combined in pairs and the microscopic HSIs paired by the two adjacent nodes are stitched, etc. Otherwise, the final two nodes in this layer are stitched first and the stitched image is set as one node in this layer, where the number of nodes is changed to even as shown in figure 8(c).

Another method is to group the images into groups of a power of two to construct several full binary trees, as shown in figure 9. If the number of input images in the synthesis tree is *N* and it can be expressed as the sum of powers of two, this synthesis tree can be transferred into multiple synthesis trees of full binary trees. At the end of each subtree synthesis, the microscopic HSIs represented by these root nodes are stitched again to obtain the final image.

In summary, a tree-based synthesis strategy has a good hierarchy, but there are large numbers of extra processing steps during the reconstruction and it does not have good universality because it relies heavily on the number of input images.

In a linear-based synthesis strategy, the two output images will not have similar size because it is a growing image synthesis strategy as shown in figure 10. Each signal block in this schematic represents a mosaic microscopic HSI and the block turns black once it is stitched until all the blocks turn black, representing the end of the mosaic. Therefore, each microscopic HSI in this method only has two kinds of status, i.e. mosaic and stitched, resulting in a simple synthesis structure. On top of this, there is no requirement for the number of input images and no specific process in this method, which means it has generality.

As a result, a linear-based synthesis strategy is adopted in this paper for the following reasons. With respect to synthesis structure, a linear-based synthesis strategy has a more straightforward and simple structure, which is easier for engineering realization. As for the input image, the tree-based synthesis
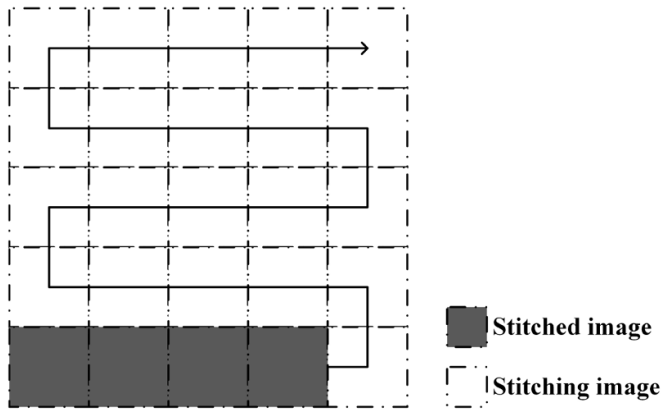
Figure 10. Schematic of linear-based synthesis strategy: captured images are stitched according to Z shape.

strategy needs to be adjusted according to the number of input images and synthesizes in a recursive way, whereas a linear-based synthesis strategy directly stitches the images without any additional processing operations. In terms of the output of the intermediate results, the two strategies will both output the microscopic HSIs at each node and store them in the storage system of the Linux server. The intermediate resultant image is the synthesis result of two images in a tree-based synthesis strategy, whereas the intermediate images in a linear-based synthesis strategy are the synthesis result of all the previous images, which can show the process of the image mosaic step by step. Regarding the output image, both synthesis strategies will output a large-scale microscopic HSI and its corresponding parameter files. To summarize, a linear-based synthesis strategy is leveraged in this paper thanks to its simple structure, friendly input and output, and good generality.

2.5. Hyperslide large-scale image specification

Large-scale microscopic HSIs will take up a huge amount of memory and hence conventional image file formats are not suitable for storing and browsing large-scale HSIs. Therefore, this paper defines a specific file format named hyperslide based on the TIFF6.0 standard for HSIs. This format can be opened with specialized software HyperViewer which reasonably adopts several development frameworks like .Net and Spring as well as various programming languages such as C#, Java, Python, IDL and Javascript. The number of pixels in length and width of the hyperslide files can be up to 2^{32} which is enough to store images of large size. However, hyperslide does not include spectral information because it consists of a hyperspectral cube false color combination pyramid image. Nevertheless, if the hyperslide file and raw file are simultaneously stored in the same folder, spectral information can be viewed in real time.

The physical structure and logical structure of a hyperslide file respectively correspond to the existence of hyperslide in the operating system and the storage structure of image data in a single hyperslide file. As shown in figure 11(a) and table 1, a hyperslide file consists of a three-stage system: image file

header (IFH), image file directory (IFD) and image data. A TIFF file contains several images and each image has a corresponding IFD. As depicted in figure 11(a), the logical structure refers to the concrete properties of IFH, IFD and directory entries (DEs), i.e. image size, data and compression mode, etc. A hyperslide file is logically composed of four groups: the triple pyramid image, thumbnail, index information and tag diagram of the sample. Among them, index information is the DE of the first-stage image and the other parts separately correspond to an IFD. The image data in hyperslide can be stored as strips and tiles. Apart from this, the triple pyramid images and thumbnails are stored in tile format, improving the efficiency of random access to large-size images.

As shown in figure 11(a), the resolution of each layer of the image in the pyramid decreases in equal proportion layer by layer and the image with the maximum resolution is the original image. The triple pyramid images in this paper contain the first stage main image and the second as well as the third stage secondary image. The resolution relationship of each layer is defined as

$$r_i = r_i \times 2^{1-i} \tag{24}$$

where r_i is the resolution of the ith-stage image.

Although the pyramid structure can increase the real-time performance of multi-resolution display, it increases the disk storage space. However, as a data cube, a single HSI can reach GB level so a stitched HSI will cause huge pressure on the storage space of ordinary computers. In addition, each single-band image of HSIs is 256 grayscale images. Human eyes can only distinguish more than 30 grayscale images, but will distinguish thousands of different shades and brightness [49]. Therefore, color images are more conducive to human eyes' interpretation than grayscale images. In order to reduce the storage burden of the system and make it easy to view images, a pseudo-color image is leveraged as the content of each stage image in this paper. The first-stage image is the pseudo-color image of the original stitched HSI array at 40× and the second-stage image is the down-sampled image of the first-stage image. Similarly, the third-stage image is obtained after secondary downsampling which is displayed on the screen.

Hyperslide is a large-scale index image file with only spatial information and no spectral information of the stitched HSIs. In order to obtain spectral information, the spatial coordinates in the large-scale HSIs can be transformed into the spatial coordinates of the specific numbered cubes in the cube array through a coordinate position mapping relationship which is demonstrated in figure 11(b). Tiles in the figure are numbered in ascending order from left to right and top to bottom and cubes are numbered in collection order. Tiles of the same color are generated from the same cube and the color also indicates exactly which part of the cube's two-dimensional space generates these tiles. JSON character strings, i.e. index information, are utilized to record the exact position of each cube in the first-stage image and they are written to the custom directory entry tag *HyperMatch* in the first-stage image.

An example of index information is shown in figure 11(c), where (*start_x, start_y*) represents the coordinate of the actual
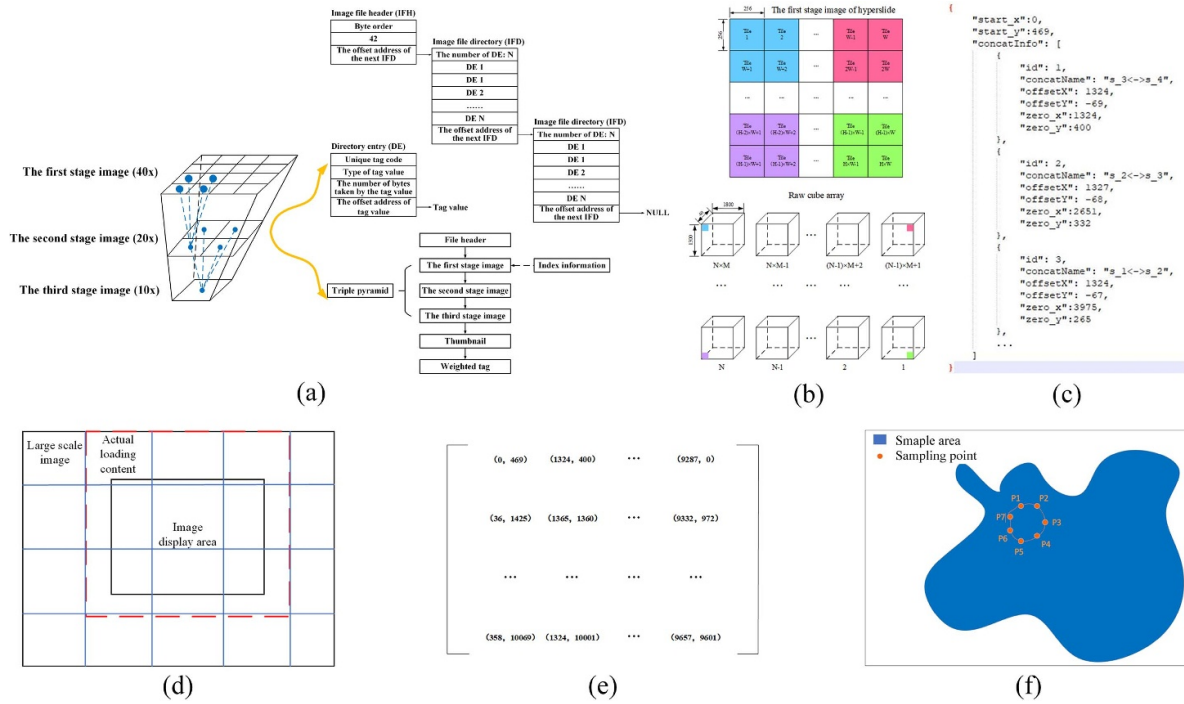
Figure 11. Hyperslide large-scale image: (a) triple pyramid image of hyperslide and its physical structure and logical structure; (b) example of the *HyperMatch* tag; (c) correspondence diagram of the first-stage image and the cube array; (d) block number to be loaded for planar dimension control (field of view controlling); (e) example graph for *HyperMatch* generating a coordinate matrix; (f) schematic diagram of the sampling point set.

Table 1. Variables in physical structure.

Variable Name	Description
IFH	IFH is located in the first eight bytes of the file, including byte order, TIFF magic word 42 and offset address of the first IFD.
IFD	IFD contains the number of DEs, several DEs and the offset address of next IFD. IFD has a data pointer DE, describing the storage address which is used for read–write interfaces to access image data.
DE	DE includes 12 bytes and is comprised of four parts: unique tag code (TIFFTag), type of tag value, the number of bytes taken by tag value and tag value or its offset address. If the number of bytes taken by tag value is less than 4, the last part stores the tag value, otherwise it stores the offset address.
TIFFTag	TIFFTag describes the inherent properties of the image, such as its width, length, compression type and offset address.

pixel starting point in the upper-left corner of the first-stage image and *concatInfo* represents the information set of the adjacent field of view generated during the mosaic process. *concatName* is composed of the numbers of two adjacent cubes (A and B), i.e. ID, and connection symbol $<->$. (*offset_x, offset_y*) is the offset vector of cube A and cube B and

(*zero_x, zero_y*) is the coordinate of the upper-left corner origin point in the two-dimensional image of cube B.

As shown in figure 11(d), due to the use of image pyramid technology, the image block with the most appropriate resolution in the pyramid will be displayed to increase the clarity. Image roaming includes three kinds of operations: up roaming, down roaming and peer roaming. In peer roaming, the feature of internal block storage is used to ensure that are only partial block data exist in the memory. Up roaming means switching from the current resolution of image to a high-resolution image, as opposed to down roaming. Peer roaming is used to operate in the same-stage image of the pyramid. Assuming that the current roaming image is in the ith layer of the pyramid, the magnification coefficient of the current display image j is from 0.25 to 1.75. If the user scales the image and $j > 1.75$, the system will default that the image of the current layer is not clear, so it will switch to the $(i-1)$th layer, i.e. up roaming, and the magnification of the ith layer will be set to 0.875. if $j < 0.25$, the system will consider that the low-resolution image is more suitable for the current display, so it will switch to the $(i+1)$th layer, i.e. down roaming, and the magnification will be set to 0.5. Otherwise, users will maintain peer roaming operations. In conclusion, switching based on this magnification scheme is insensitive, smooth and infrequent.

The spectral curve of a hyperslide file can be obtained from the corresponding cube based on the mapping relationship. First, it is shown in figure 11(e) that a coordinate matrix is established according to (*start_x, start_y*) and (*zero_x, zero_y*) in *HyperMatch*. The element in the matrix represents the upper-left corner origin point (0, 0) of the two-dimensional

index of each cube in the first-stage image. The operation of the viewing spectrum can be divided into two types: viewing a single-point spectrum and viewing a regional spectrum. Viewing a single-point spectrum can be modeled as a two-dimensional plane search problem and viewing a regional spectrum can subsequently be simplified as multiple single-point search problems.

2.5.1. Single-point spectral mapping algorithm   Single-point spectral mapping is triggered when the point (x, y) in the image is clicked. Assuming that the coordinates of the origin point of the cube (two-dimensional size is H $\times$ W) in the first-stage image are (x', y'), if (x, y) satisfies (25), it means that this point is located within the cube and its absolute coordinate is $(x - x', y - y')$. Then, the spectral information $S(x - x', y - y')$ can be obtained through the absolute coordinate. However, the time complexity of this solution is $O(mn)$ which is not efficient. Therefore, the order of the matrix and variability of search step length are used to reduce the search time and improve the real-time performance of the software.

$$
\begin{cases}
x' \leqslant x \leqslant x' + W \\
y' \leqslant y \leqslant y' + H
\end{cases}.
\tag{25}
$$

It is obvious in figure 11(e) that the x-coordinates of pixels in each row increase from left to right and top to bottom, and the y-coordinates of pixels in each column increase from top to bottom and right to left. According to the order of the matrix, the binary search method will speed up the algorithm and the time complexity will be $O(\lg(mn))$. However, this method cannot satisfy the variability of search step length. Therefore, an heuristic algorithm is designed in this paper: look for the position in the horizontal direction first and the position in the vertical direction second. Then, the upper-right corner of the matrix is selected as the search starting point by setting the direction marker *DTrigger* to control the search direction (1 means horizontal, -1 means vertical). The specific search steps are as follows:

If $x \leq x'$, search to the left until $x' \leq x \leq x' + W$; If $x' + W \leq x$, search to the bottom until $x' \leq x \leq x' + W$. Then set *DTrigger* = *-DTrigger*.

If $y' + W \leq y \leq y' + H$, set *DTrigger* = *-DTrigger* and reevaluate whether x is within the range. If the condition is satisfied, the cube to which the coordinate belongs can be found out. Otherwise, repeat step a. If $y \leq y'$, search to the up until $y' \leq y \leq y' + H$; If $y' + H \leq y$, search to the bottom until $y' \leq y \leq y' + H$. Then, repeat step b.

This method does not need to traverse the entire two-dimensional matrix to find the target. Simultaneously, in the case of constant step size, the time complexity is $O(m + n)$ and the length of a search path is the Manhattan distance (the shortest).

Additionally, the variability of the search step length is also useful for improving search efficiency. The variables used in this method are shown in table 2.

$\triangle x$ and $\triangle y$ are defined as

$$
\begin{cases}
\Delta x = |x - x'| \\
\Delta y = |y - y'|
\end{cases}.
\tag{26}
$$

Table 2. Variables used in search.

Variable name	Variable symbol
Unit step length	Step_{min}
Lateral displacement difference threshold	Threshold_{row}
Longitudinal displacement difference threshold	Threshold_{col}
Horizontal distance	$\triangle x$
Vertical distance	$\triangle y$

The step size for each move is defined as

$$
\begin{cases}
\text{Step}_x = \Delta x / \text{Threshold}_{row} \cdot \text{Step}_{min} \\
\text{Step}_y = \Delta y / \text{Threshold}_{col} \cdot \text{Step}_{min}
\end{cases}.
\tag{27}
$$

The default unit step size is set as 1, i.e. one grid at a time, and the empirical values for Threshold_{row} and Threshold_{col} are 1350 and 950 (depending on the cube size). With the help of variable search step size, the search distance is shorter and thus the number of attempts is greatly reduced and the time complexity can be reduced to $O(1)$.

2.5.2. Regional spectral mapping algorithm   The regional spectral mapping algorithm can be triggered by drawing a closed area or curve. This problem can then be simplified to multiple single-point spectral mapping by sampling M coordinate points of the circle trace or the closed region, i.e. $P_1, P_2, \dots, P_M$. The regional spectra can be obtained by adding and averaging the spectra at all points of the point set which are acquired through single-point spectral mapping. This method is quite simple but ignores the spatial correlation of each point and always needs to search from the top right corner of the two-dimensional array. Therefore, the search time will be too long if M and the two-dimensional array are too large. Consequently, after the mapping of the initial sampling point P_1 is completed, the search starting point of the remaining point P_n is set to the cube which contains P_{n-1} (P_{n-1} and P_n are adjacent as shown in figure 11(f)). This method can effectively shorten the search time. If P_{n-1} and P_n do not belong to the same cube, the cube number of P_n will be ideally found by moving at most one step.

3. Experimental results and analysis

All data used in this paper were collected by a homemade AOTF microscopic hyperspectral imaging system. The HSI system used in this paper is composed of a microscope (Nikon 80i, Nikon Corp.), an AOTF (VA310-.37-.80-L, Brimrose Corp.), an AOTF controller (VFI130-140SPFB2C2exSTS, Brimrose Corp.), a gray scientific complementary metal oxide semiconductor (sCMOS, Dhyana 400D, Tucsen Corp.), a color charge-coupled device detector (color CCD, DigiRetina 16, Tucsen Corp.) and a personal computer [7]. In addition, multidimensional imaging software is developed to control these devices, supplying a friendly interface to obtain color images and HSIs simultaneously. On top of this, the autofocus

Figure 12. Physical map of the CW-DY objective micrometer. It has three standard scales: 10 mm subdivided into 100 cells, 1 mm subdivided into 100 cells and 10 mm subdivided into 200 cells.
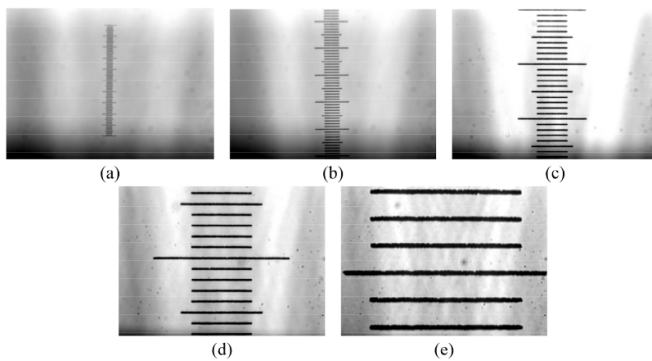

(a)　　　　　　(b)　　　　　　(c)

(d)　　　　　　(e)

Figure 13. Images captured under different objective multiples: (a) 4×; (b) 10×; (c) 20×; (d) 40×; (e) 100×.

algorithm in this system made it possible to acquire large-scale microscopic HSIs and color images.

3.1. Measurement of magnification factor

In this system, the relative position of two adjacent microscopic HSIs is used to correct the offset vectors. However, the relative position is based on the physical coordinate system of the platform, while the vector correction is based on the pixel offset. Therefore, a magnification factor is needed for the conversion between the two quantities and the magnification factor is related to the microscope objective multiple. The measurement of the magnification factor is operated by capturing the objective micrometer, which is a calibration facility of size 785 mm × 25 mm. The objective micrometer used in this system is a CW-DY objective micrometer from Beijing Pediwei Instrument Co. Ltd. It is shown in figure 12 that the objective micrometer has three standard scales: 10 mm subdivided into 100 cells, 1 mm subdivided into 100 cells and 10 mm subdivided into 200 cells.

The images of the second standard scale are captured under different objective multiples in figure 13 and the number of pixels per unit calibration is calculated from these images. As illustrated in table 3, several different calibrations are captured and they are averaged in consideration of image quality and artificial measurement error. These magnification factors are utilized in offset correction.

3.2. Image registration and refusion

It is shown in figure 14 that there is some lateral offset between the two adjacent mosaic microscopic HSIs. The two images are stitched through the SURF method ($\vec{p}_{\text{init}}$), position vector calculation ($\vec{p}_{\text{pos}}$) and the image fusion method ($\vec{p}_{\text{final}}$). To verify the necessity of the three processing steps, three experiments are operated in this paper. First, only the SURF method is used in linear-based synthesis and the result in figure 15(a) shows that there will be displacement in the stitched image. Second, only the triaxial coordinate information is used in linear-based synthesis and the result is similar to the first experiment shown in figure 15(b). Third, image fusion is adopted in linear-based synthesis, generating the correct resultant image in figure 15(c).

If the microscopic HSI is stitched only using its texture information, there will be offset in the stitched image. In addition, the coordinate-information-based matching method will import movement error which will be magnified to a large image offset in the microscopic imaging system. Although both types of information will lead to some offset, they are indispensable in an image mosaic. Therefore, the microscopic HSI mosaic method based on the spatial redundancy characteristic and position correction integrates the basic image matching algorithm, the position information of the imaging system and the HSI information, effectively improving the accuracy of the image mosaic. Besides, the multi-wavelength fusion method is adopted in this paper, removing wavelength selection and enhancing the efficiency of the algorithm. Compared with existing fusion methods, there is no need for both HSIs and MSIs in the multi-wavelength fusion method, avoiding spending too much time on image collection.

3.3. Blank microscopic HSI screening

Continuous multiple histopathological HSIs with large blank areas will seriously influence the autofocus and large-scale microscopic HSI mosaic. Therefore, this experiment calculates the ratio of blank area in the histopathological HSIs. First, one microscopic HSI with histopathological tissue and one blank image are taken in the system and then the ratio image of the two images calculated by (20) is shown in figure 16(a), i.e. preprocessed image. To calculate the ratio of blank area in this image, the image is divided into two parts with the average brightness of the image as the threshold. Then, the ratio image is turned into a binary image with the same threshold, as illustrated in figure 16(c). Next, the ratio of blank area is separately calculated in six parts of the binary image in figure 16(d). If the ratio is higher than a certain threshold, the autofocus will ignore the concerned HSI in the case of blurred HSIs and then the sharpness of the stitched image will be ensured.

This method can also be utilized to compare two blank microscopic HSIs to screen a clearer blank image for image preprocessing. The screening will supply a better blank image as the reference image in autofocus and assist in image preprocessing with better spectral correction and removal of intrinsic interference in the optical path. Given the two blank microscopic HSIs in figure 17, it is obvious that there is less dust in

Table 3. Magnification factor between the number of pixels and unit calibration.

Objective	Unit Calibration (mm)					Magnification Factor
	0.01	0.02	0.05	0.1	0.2	
4×	9.513	17.846	44.698	89.045	177.113	902.715
10×	23.371	45.044	110.432	220.531	442.566	2243.216
20×	46.043	91.611	222.090	448.286	891.022	4521.924
40×	87.052	178.817	443.807	896.807	–	8872.565
100×	222.081	453.486	453.486	–	–	22311.7
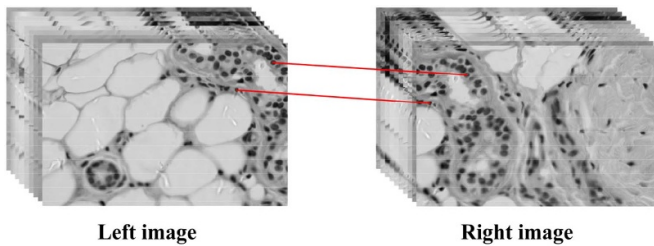
Figure 14. Schematic map of two adjacent microscopic HSIs.

Figure 15. Resultant image of different methods: (a) only using the SURF method; (b) only using triaxial coordinate information; (c) using image fusion.

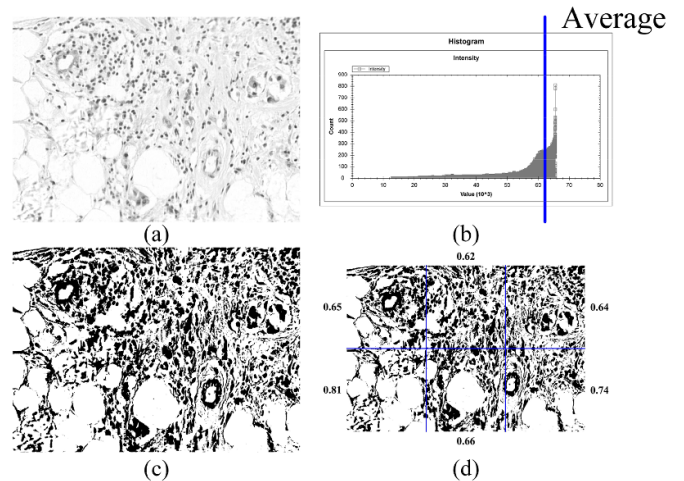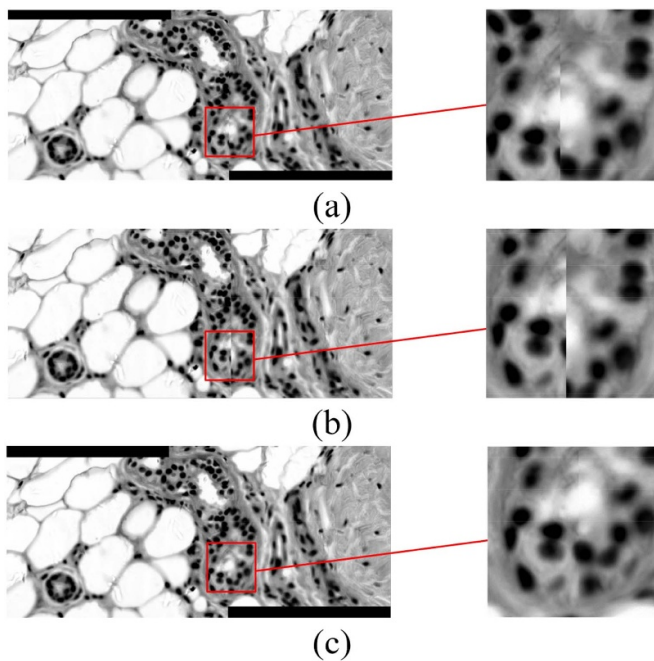**Figure 16.** (a) Ratio image; (b) histogram of the image; (c) binary image; (d) ratio of blank area in six parts.

Figure 17. Respective ratios of blank area in the two blank microscopic HSIs.

image (b), so it is set as the clearer blank image in calculating the ratio of blank area. Simultaneously, it is calculated that the ratio of blank area in image (a) is 3.83 and that in image (b) is 4.01, meaning that image (b) is clearer. A unified conclusion is drawn from both subjective and objective aspects, verifying the reliability of the algorithm.
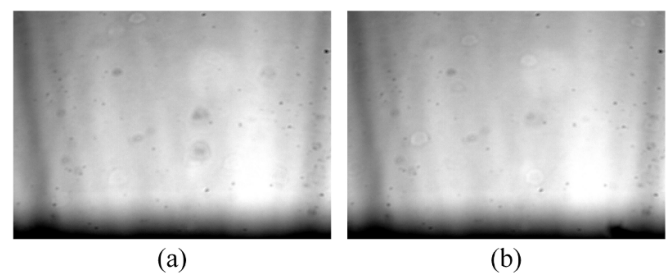
3.4. Large-scale microscopic HSI generation

The linear-based synthesis strategy is leveraged to generate a large-scale multidimensional image. This paper displays the stitched image of four images of a melanoma sample as well as the intermediate output images. In this experiment, four HSIs of a melanoma sample (40×) are collected with 30 bands from 400 nm to 800 nm. Figure 18 shows the 15th band, i.e. 606 nm, of the stitched image of four microscopic HSIs. It is demonstrated that the multidimensional image mosaic method used in this paper realizes large-scale pathological HSI acquisition in figure 19(d) and solves the problem of small field of view at high magnification. The resultant HSI contains both rich information and a large observation field which
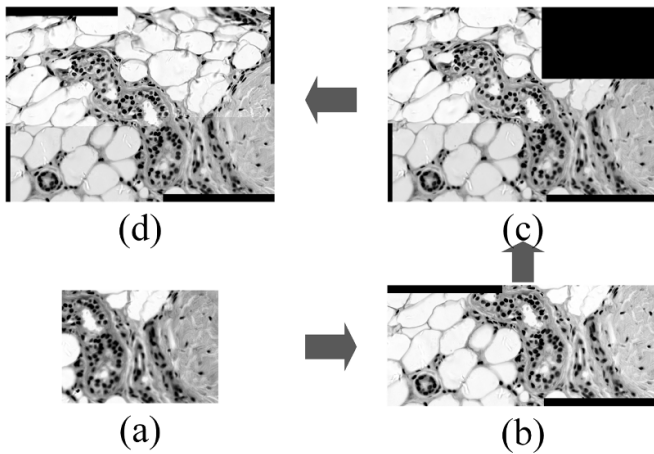
Figure 18. Mosaic process of four microscopic HSIs in melanoma samples at 40×. (a) Lower right corner of the final stitched image; (b) stitched image of two images; (c) stitched image of three images; (d) stitched image of four images.
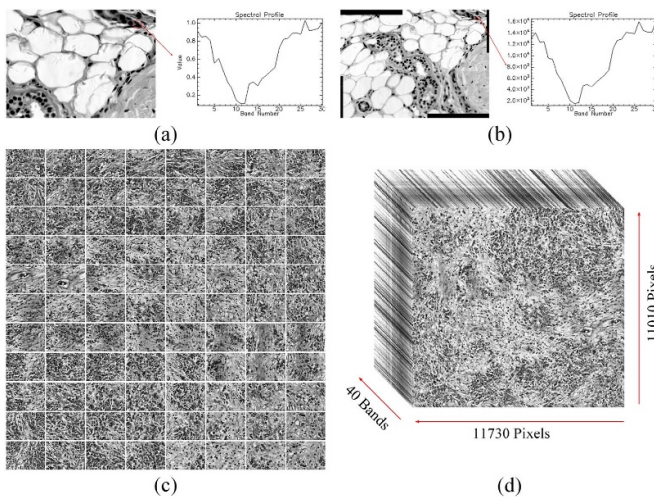

Figure 19. (a) 15th band of pre-stitched HSI and spectral curve; (b) 15th band of stitched HSI and spectral; (c) 88 15th-band images with single field of view; (d) large-scale microscopic HSI of a lymph node sample taken at 40×.

Figure 20. Software interface of HyperViewer. The image shown on the screen is the pseudo-color image of the original HSI.

can be browsed in figure 20, providing more extensive and in-depth information for the subsequent pathological analysis. As

Table 4. Main elements in HyperViewer.

Variable Name	Description
(a)	**Pseudo-color image display area:** The displayed image is the pseudo-color image of the large-scale microscopic HSI.
(b)	**Toolbar:** The image dragging tool, annotation tool, spectrum acquisition tool and history back tool are included here.
(c)	**Spectrum curve rendering area:** When the spectrum acquisition tool is clicked, the spectrum is displayed in this area. The horizontal axis is the wavelength and the vertical axis is the transmittance or pixel value, and they are adaptive to the band and pixel value range of the underlying cube.
(d)	**Thumbnail interactive area:** In addition to dragging the image in the central image display area, users can also roam in the thumbnail interactive area. The yellow box represents the relative position of the current display area in a large-scale image. By moving the yellow box, fast positioning can be realized under a high-multiple field of view.
(e)	**Snapshot of regions of interest:** When a user is interested in part of the image, the image interception function can be used to intercept the two-dimensional index and then the composite cube can be exported according to the two-dimensional index.
(f)	**Magnification dial:** The magnification button on the magnification dial can be used to zoom in and out of the central image area quickly.
(g)	**Section label area:** This area displays the label information of the pathological section entity corresponding to the current image.
(h)	**Annotation information display area:** Annotation information will be displayed in this area after circling in the displayed image with the annotation tool.

shown in figure 20, the initial display multiple is 2.5× and the main elements in HyperViewer are introduced in table 4. Moreover, as shown in figures 19(a) and (b), the shape of the spectral curves of the pre-stitched HSI and stitched HSI are almost the same, which reveals that the mosaic method will preserve the spectral characteristics of HSIs.

To evaluate the proposed method,

Table 5 compares SURF with SIFT on ten microscopic HSIs of lymph node samples in the feature detection time. The comparison experiments are operated on a 2.3 GHz Intel Core i5 GPU with 16 GB 2133 MHz LPDDR3 memory. The system environment is MacOS Mojave 10.14.4. It is obvious that SURF outperforms SIFT in all experiments. SURF can save half the detection time in comparison to SIFT.

Moreover, the RMSE (root-mean-square error) and CS (cosine similarity) are leveraged to evaluate the accuracy of matching results. RMSE is a standard statistical metric to

Table 5. Time consumption comparison of feature point detection.

Number	SIFT (ms)	SURF (ms)
1	760.22	233.88
2	501.55	211.37
3	732.70	224.19
4	642.66	218.28
5	471.46	227.74
6	510.27	216.61
7	512.08	231.43
8	460.40	226.85
9	449.59	199.29
10	476.71	207.84

Table 6. Comparison of matching precision in the first nine matches.

	SIFT + RANSAC		SURF + RANSAC	
Number	RMSE	CS	RMSE	CS
1	4.33	0.9969	4.17	0.9969
2	3.12	0.9949	53.06	0.9933
3	3.04	0.9977	3.24	0.9977
4	3.28	0.9980	2.73	0.9980
5	4.12	0.9970	3.72	0.9972
6	709.16	0.9853	580.15	0.9921
7	3.53	0.9967	4.08	0.9970
8	2.77	0.9986	2.98	0.9986
9	4.11	0.9959	4.04	0.9964

measure model performance, calculating the coordinate error between matching points of two images [50]. It is given by

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(H_{12}X_i - Y_i)^2} \qquad (28)$$

where X_i and Y_i represent the interior point coordinates of two microscopic HSIs (I_1, I_2) which are to be stitched together. N represents the number of interior points. H_{12} represents the transformational matrix from I_1 to I_2. The smaller the RMSE, the more accurate the match.

CS measures the spectral angle between two vectors to calculate the spectral similarity of two images [51]. It is given by

$$\overline{\cos\langle X, Y\rangle} = \frac{1}{N}\sum_{i=1}^{N}\frac{x_i^T y_i}{\sqrt{x_i^T x_i}\sqrt{y_i^T y_i}} \qquad (29)$$

where x_i and y_i represent the spectral curves of two matching feature points (X and Y). The smaller the angle between two vectors, the more similar the two vectors are and the more accurate the algorithm is. Table 6 compares the matching accuracy of the first nine matches in the lymph node image mosaic. The RMSE and CS of the two matching algorithms are close and their CE are both very high. However, SURF can detect feature points much faster than SIFT.

4. Conclusions

In this work, we have presented a microscopic HSI mosaic method to generate large-scale microscopic HSIs. The problem is simplified into the mosaic of two adjacent images, so the images will be stitched well when the offset vectors of the two images are calculated precisely. This paper utilizes a SURF-based mosaic method to extract the features of adjacent HSIs and meanwhile the position information is used to assist in obtaining more accurate offset vectors. Then, a spatial redundancy-based multi-wavelength image matching method is applied in this paper thanks to different texture structures and grayscale between different wavelengths to obtain the final fusion vectors. Moreover, a blank microscopic HSI screening method is also introduced to obtain a clearer blank image for better image preprocessing. Finally, adjacent microscopic HSIs are merged by the linear-based synthesis method due to its simple synthesis structure and good generality. The method in this paper will save a lot of time in image acquisition because the fusion method is only based on HSIs instead of both HSIs and MSIs. There is no need for special instruments in image acquisition or particular sample types in the image mosaic. The experiments demonstrate the necessity of combining SURF and coordinate information and there are no gaps in the resultant images of the microscopic HSI mosaic method, proving its validity and practicability. In addition, this paper defines a file format that is named hyperslide and opened with software HyperViewer for large-scale microscopic HSI browsing. However, each group of microscopic HSIs is preprocessed by the same blank image, ignoring the possible deviations of the incident light. Moreover, the proposed method relies on the imaging technology and mosaic algorithms of two-dimensional images and there may be some breakthrough in future research.

ORCID iD

Qingli Li ⓘ https://orcid.org/0000-0001-5063-8801

References

[1] Schneckenburger H *et al* 1994 Time-gated microscopic imaging and spectroscopy in medical diagnosis and photobiology [also Erratum 33(11)3828(Nov1994)] *Opt. Eng.* **33** 2600–6
[2] Morey A F, Bruce J E and McAninch J W 1996 Efficacy of radiographic imaging in pediatric blunt renal trauma *J. Urol.* **156** 2014–8

[3] Delsanto S, Molinari F, Giustetto P, Liboni W, Badalamenti S and Suri J S 2007 Characterization of a completely user-independent algorithm for carotid artery segmentation in 2-D ultrasound images *IEEE Trans. Instrum. Meas.* **56** 1265–74

[4] Duchateau N, Sermesant M, Delingette H and Ayache N 2018 Model-based generation of large databases of cardiac images: synthesis of pathological cine MR sequences from real healthy cases *IEEE Trans. Med. Imaging* **37** 755–66

[5] Gutiérrez Castorena E V, Gutiérrez-Castorena M D C, González Vargas T, Cajuste Bontemps L, Delgadillo Martínez J, Suástegui Méndez E and Ortiz Solorio C A 2016 Micromapping of microbial hotspots and biofilms from different crops using digital image mosaics of soil thin sections *Geoderma* **279** 11–21

[6] Thévenaz P and Unser M 2007 User-friendly semiautomated assembly of accurate image mosaics in microscopy *Microsc. Res. Tech.* **70** 135–46

[7] Zhang Q, Li Q, Yu G, Sun L, Zhou M and Chu J 2019 A multidimensional choledoch database and benchmarks for cholangiocarcinoma diagnosis *IEEE Access* **7** 149414–21

[8] Li L, Xu M, Liu H, Li Y, Wang X, Jiang L, Wang Z, Fan X and Wang N 2020 A large-scale database and a CNN model for attention-based glaucoma detection *IEEE Trans. Med. Imaging* **39** 413–24

[9] Hosseini M S, Brawley-Hayes J A Z, Zhang Y, Chan L, Plataniotis K and Damaskinos S 2020 Focus quality assessment of high-throughput whole slide imaging in digital pathology *IEEE Trans. Med. Imaging* **39** 62–74

[10] Mukhopadhyay S *et al* 2018 Whole slide imaging versus microscopy for primary diagnosis in surgical pathology: a multicenter blinded randomized noninferiority study of 1992 cases (pivotal study) *Am. J. Surg. Pathol.* **42** 39–52

[11] Pantanowitz L, Sinard J H, Henricks W H, Fatheree L A, Carter A B, Contis L, Beckwith B A, Evans A J, Lal A and Parwani A V 2013 Validating whole slide imaging for diagnostic purposes in pathology: guideline from the College of American Pathologists Pathology and Laboratory Quality Center *Arch. Pathol. Lab. Med.* **137** 1710–22

[12] Kawano Y, Namiki K, Miyawaki A and Ishikawa T 2017 Extending whole slide imaging: color darkfield internal reflection illumination (DIRI) for biological applications *PLoS One* **12** e0167774

[13] Guo C F, Bian Z, Jiang S, Murphy M, Zhu J, Wang R, Song P, Shao X, Zhang Y and Zheng G 2020 OpenWSI: a low-cost, high-throughput whole slide imaging system via single-frame autofocusing and open-source hardware *Opt. Lett.* **45** 260–3

[14] Wang Z and Yang Z 2020 Review on image-stitching techniques *Multimedia Syst.* **26** 1–18

[15] Shan Y 2014 The study on auto mosaic technology of microscopic image *Int. Conf. on Communication Systems and Network Technologies* (*Bhopal, India*) pp 843–6

[16] Pang M, Chen Z F and Liu S 2019 Scale invariant feature transform-Gauss mosaic method for blood cell microscopic images *J. Med. Imaging Health Inf.* **9** 134–9

[17] Li Q, Zhou M, Liu H, Wang Y and Guo F 2015 Red blood cell count automation using microscopic hyperspectral imaging technology *Appl. Spectrosc.* **69** 1372–80

[18] Wang Q, Wang J, Zhou M, Li Q and Wang Y 2017 Spectral-spatial feature-based neural network method for acute lymphoblastic leukemia cell identification via microscopic hyperspectral imaging technology *Biomed. Opt. Express* **8** 3017–28

[19] Brückner M, Becker K, Popp J and Frosch T 2015 Fiber array based hyperspectral Raman imaging for chemical selective analysis of malaria-infected red blood cells *Anal. Chim. Acta* **894** 76–84

[20] Liao J, Jiang S, Zhang Z, Guo K, Bian Z, Jiang Y, Zhong J and Zheng G 2018 Terapixel hyperspectral whole-slide imaging via slit-array detection and projection *J. Biomed. Opt.* **23** 066503–066503

[21] Pian Q, Yao R, Sinsuebphon N and Intes X 2017 Compressive hyperspectral time-resolved wide-field fluorescence lifetime imaging *Nat. Photon.* **11** 411–4

[22] Yang X, Kwitt R, Styner M and Niethammer M 2017 Quicksilver: fast predictive image registration – A deep learning approach *NeuroImage* **158** 378–96

[23] Klein S, Staring M, Murphy K, Viergever M A and Pluim J 2010 elastix: a toolbox for intensity-based medical image registration *IEEE Trans. Med. Imaging* **29** 196–205

[24] Yu D, Yang F, Yang C, Leng C, Cao J, Wang Y and Tian J 2016 Fast rotation-free feature-based image registration using improved N-SIFT and GMM-based parallel optimization *IEEE Trans. Biomed. Eng.* **63** 1653–64

[25] Gong M, Zhao S, Jiao L, Tian D and Wang S 2014 A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information *IEEE Trans. Geosci. Remote Sens.* **52** 4328–38

[26] Sakai Y *et al* 2015 An object tracking system based on SIFT and SURF feature extraction methods *Int. Conf. on Network-Based Information Systems* (*Taipei, Taiwan*) pp 561–5

[27] Lowe D G 2004 Distinctive image features from scale-invariant keypoints *Int. J. Comput. Vis.* **60** 91–110

[28] Wang V T and Hayes M P 2017 Synthetic aperture sonar track registration using SIFT image correspondences *IEEE J. Ocean. Eng.* **42** 901–13

[29] Al-khafaji S L, Zhou J, Zia A and Liew A W-C 2018 Spectral-spatial scale invariant feature transform for hyperspectral images *IEEE Trans. Image Process.* **27** 837–50

[30] Sedaghat A and Mohammadi N 2019 High-resolution image registration based on improved SURF detector and localized GTM *Int. J. Remote Sens.* **40** 2576–601

[31] Song Z L and Zhang J 2010 Remote sensing image registration based on retrofitted SURF algorithm and trajectories generated from lissajous figures *IEEE Geosci. Remote Sens. Lett.* **7** 491–5

[32] Bay H and Tuytelaars T 2006 *SURF: Speeded Up Robust Features* (Berlin: Springer) pp 404–17

[33] Cetin M and Tepecik A 2016 Intensity-hue-saturation-based image fusion using iterative linear regression *J. Appl. Remote Sens.* **10** 045019–045019

[34] Yang B and Li S 2010 Multifocus image fusion and restoration with sparse representation *IEEE Trans. Instrum. Meas.* **59** 884–92

[35] Wang Q, Shen Y I, Zhang Y E and Qiu Zhang J 2003 A quantitative method for evaluating the performances of hyperspectral image fusion *IEEE Trans. Instrum. Meas.* **52** 1041–7

[36] Wang Q, Shen Y, Zhang Y and Zhang J Q 2004 Fast quantitative correlation analysis and information deviation analysis for evaluating the performances of image fusion techniques *IEEE Trans. Instrum. Meas.* **53** 1441–7

[37] Fu H, Bian L, Cao X and Zhang J 2020 Hyperspectral imaging from a raw mosaic image with end-to-end learning *Opt. Express* **28** 314–24

[38] Zhang K, Wang M, Yang S and Jiao L 2018 Spatial-spectral-graph-regularized low-rank tensor decomposition for multispectral and hyperspectral image fusion *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **11** 1030–40

[39] Sylla D, Minghelli-Roman A, Blanc P, Mangin A and Hembise Fanton d'Andon O 2014 Fusion of multispectral

images by extension of the pan-sharpening ARSIS method *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **7** 1781–91

[40] Chen Z, Pu H, Wang B and Jiang G-M 2014 Fusion of hyperspectral and multispectral images: a novel framework based on generalization of pan-sharpening methods *IEEE Geosci. Remote Sens. Lett.* **11** 1418–22

[41] Palsson F, Sveinsson J R, Ulfarsson M O and Benediktsson J A 2015 Model-based fusion of multi- and hyperspectral images using PCA and wavelets *IEEE Trans. Geosci. Remote Sens.* **53** 2652–63

[42] Zhou Y, Feng L, Hou C and Kung S-Y 2017 Hyperspectral and multispectral image fusion based on local low rank and coupled spectral unmixing *IEEE Trans. Geosci. Remote Sens.* **55** 5997–6009

[43] Abdlaty R, Orepoulos J, Sinclair P, Berman R and Fang Q 2018 High throughput AOTF hyperspectral imager for randomly polarized light *Photonics* **5** 3

[44] Abdlaty R, Wong R K W, Hayward J E, Farrell T J and Fang Q 2018 Hyperspectral imaging and classification for grading skin erythema *Front. Phys.* **6** 1–10

[45] Ma L, Crawford M M and Tian J 2010 Local manifold learning-based k-nearest-neighbor for hyperspectral image classification *IEEE Trans. Geosci. Remote Sens.* **48** 4099–109

[46] Zhang Y, Lu S, Zhou X, Yang M, Wu L, Liu B, Phillips P and Wang S 2016 Comparison of machine learning methods for stationary wavelet entropy-based multiple sclerosis detection: decision tree, k-nearest neighbors, and support vector machine *Simulation* **92** 861–71

[47] Raguram R, Chum O, Pollefeys M, Matas J and Frahm J-M 2013 USAC: a universal framework for random sample consensus *IEEE Trans. Pattern Anal. Mach. Intell.* **35** 2022–38

[48] Jeon S, Song H B, Kim J, Lee B J, Managuli R, Kim J H, Kim J H and Kim C 2017 In vivo photoacoustic imaging of anterior ocular vasculature: a random sample consensus approach *Sci. Rep.* **7** 4318–9

[49] Jiang H, Zhao Y, Gao J and Gao Z 2017 Adaptive pseudo-color enhancement method of weld radiographic images based on HSI color space and self-transformation of pixels *Rev. Sci. Instrum.* **88** 065106

[50] Chai T and Draxler R R 2014 Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature *Geosci. Model Dev.* **7** 1247–50

[51] Han S *et al* 2015 Cosine similarity based fingerprinting algorithm in WLAN indoor positioning against device diversity *IEEE Int. Conf. on Communications* (*London, UK*) pp 2710–4