MSIr: Automatic Registration Service for Mass Spectrometry Imaging and Histology

Bo-Jhang Lin, Tien-Chueh Kuo, Hsin-Hsiang Chung, Ying-Chen Huang, Ming-Yang Wang, Cheng-Chih Hsu, Po-Yang Yao, and Yufeng Jane Tseng*

Cite This: *Anal. Chem.* 2023, 95, 3317−3324

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Mass spectrometry imaging (MSI) is a powerful tool that can be used to simultaneously investigate the spatial distribution of different molecules in samples. However, it is difficult to comprehensively analyze complex biological systems with only a single analytical technique due to different analytical properties and application limitations. Therefore, many analytical methods have been combined to extend data interpretation, evaluate data credibility, and facilitate data mining to explore important temporal and spatial relationships in biological systems. Image registration is an initial and critical step for multimodal imaging data fusion. However, the image registration of multimodal images is not a simple task. The property difference between each data modality may include spatial resolution, image characteristics, or both. The image registrations between MSI and different imaging techniques are often achieved indirectly through histology. Many methods exist for image registration between MSI data and histological images. However, most of them are manual or semiautomatic and have their prerequisites. Here, we built MSI Registrar (MSIr), a web service for automatic registration between MSI and histology. It can help to reduce subjectivity and processing time efficiently. MSIr provides an interface for manually selecting region of interests from histological images; the user selects regions of interest to extract the corresponding spectrum indices in MSI data. In the performance evaluation, MSIr can quickly map MSI data to histological images and help pinpoint molecular components at specific locations in tissues. Most registrations were adequate and were without excessive shifts. MSIr is freely available at https://msir.cmdm.tw and https://github.com/CMDM-Lab/MSIr.
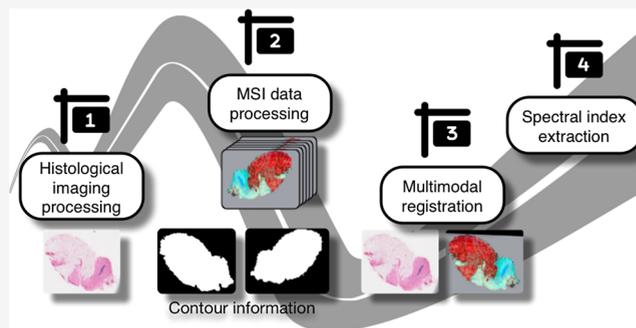
INTRODUCTION

Mass spectrometry imaging (MSI) is a powerful, label-free, and increasingly used analytical tool for investigating the spatial distribution of molecules. To acquire MSI, the sample surface is analyzed pixel-by-pixel with a mass spectrometer. By selecting a peak in the resulting spectrum that corresponds to the compound of interest, the MS data can be used to map its distribution in the sample. The spatial distribution of multiple molecules in specific molecular classes, such as proteins, peptides, metabolites, lipids, and glycans, can be analyzed through MSI in a single experiment.[1] Thus, MSI has been exploited in many fields, such as proteomics,[2] pharmaceutical research and development,[3] cancer diagnostics,[4] forensics,[5] and natural product analysis.[6]

Each analytical technique has different usage limitations and analytical properties.[7] An increasing number of analytical techniques are being used together. Similarly, to analyze more metabolites and improve data interpretation, MSI with multiple ionization techniques is increasingly being used.[1] In addition, different imaging techniques from radiology,[1] spectroscopy,[7] and microscopy[1,7] have been combined with MSI.

Magnetic resonance imaging (MRI) is a medical imaging technique offering 3D anatomical structure information of a sample with high spatial resolution and a precise sample shape. However, MRI cannot be used to monitor metabolic information. Thus, MRI combined with MSI may be a powerful tool to connect human diseases and metabolic disorders through biomolecular pathways. For example, desorption electrospray ionization MSI was used in an oncometabolite to guide brain tumor surgery, and the distribution of the target metabolite was mapped to 3D MRI to assist clinical decision-making.[8] Vibrational spectroscopy imaging (VSI) is a technique to investigate the distribution and composition of molecules from a sample in a label-free manner. In VSI methods, the vibrations of chemical bonds, which

provide the fingerprint signature of a molecule, are examined. Different vibrational modes are often used to investigate and identify specific chemical bonds in a molecule. VSI can be used to examine the abundance, structure, and conformation of biomolecules, complementing the problem that MSI cannot be used to identify isomers, enantiomers, isobars, and neutral molecules.[9] Microscopy provides high-resolution morphological and structural information that is lacking in MSI. MSI and microscopy are currently the most common multimodal systems used to identify and classify tissue sections. In bright-field microscopy, histological stains are often used to highlight tissue morphology and cell structure with a specific dye; in cancer or functional disorder research, tissues have been routinely stained after MSI experiments to study the molecular profile relationship with the histological features of normal or abnormal tissue regions.[10]

Image registration is the initial and necessary step in most multimodal imaging data fusion methods. However, each modality's spatial resolution and image characteristics vary, which makes the registration of multimodal images challenging.[7] Several methods can help register histology and MSI, but most have prerequisites. One method is image registration through matching reference points. MSI and histological images are registered by aligning these reference points using linear transformations (rigid or affine).[7] Another method is semiautomatic registration between MSI data and histological matches through autofluorescence microscopy images.[11] In this method, manually matching the laser ablation markers and MS image pixels was needed, and the method was only suitable for ionization techniques that generate laser ablation markers.

Data-driven registration is a method that has recently gained more attention. In brief, MSI data are first summarized by multivariate techniques to a single representative image. Then, an image property of the representative image, such as the tissue region, contour, or histoanatomical structure, is used to register histology. To date, several automatic data-driven registration methods and tools have been proposed. Veselkov et al. developed a comprehensive framework including automatic registration for MSI data analysis.[12] In this image registration approach, affine transformations were optimized using the gradient descent optimization method to maximize the number of overlapped tissue-related pixels between the MS image and another modal image. Abdelmoula et al. developed a method for the automatic generic registration of MSI data and histological data via the nonlinear dimensionality reduction method t-distributed stochastic neighbor embedding (t-SNE).[13] The t-SNE method is used to summarize the MSI data and project them onto a nonlinear data structure while preserving their features. It provides sufficient histoanatomical structure information to register MSI data to histological images of the same or adjacent regions. Patterson et al. proposed the R package RegCombIMS,[14] an extension of the R package Cardinal,[15] which provides automatic registration between multiple MSI data sets targeting different analytes through histological image intermediates.

Each method has its prerequisites, such as labeling reference points or an autofluorescence microscope, and not every MSI experiment can satisfy them. Among current image registration methods, the method proposed by Abdelmoula et al.[13] is generic under any situation, but the t-SNE algorithm is not scalable and time-consuming. Recently, Race et al. proposed a modified method[16] to solve the speed problem of t-SNE. However, there is no ready-made software to assist in image

registration between MSI data and histological images for users without an image processing background or coding skills. Thus, we provide MSI Registrar (MSIr), an easy-to-use free web service for automatic image registration between MSI data and the corresponding histological image.

■ RESULTS AND DISCUSSION

Overview of MSIr. MSIr is a web-based platform with Node.js (version 14.16.0) and the Express framework (version 4.17.1). The user interfaces of MSIr are built through the React library (version 17.0.2). MSI data and histological image registration was implemented using Python (version 3.8.2). MSI data must be converted to centroid imzML format before execution. The main procedures in MSIr include histological image processing, MSI data processing, multimodal registration, and spectral index extraction, as shown in Figure 1 (please see Figure S1 for details).
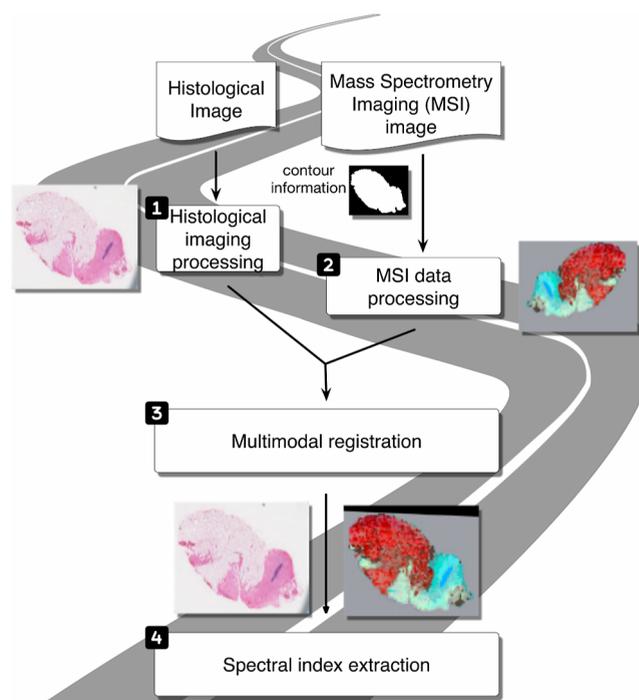

Figure 1. Procedures in the MSIr.

Histological Image Processing. Our method was used to successfully separate the tissue regions and backgrounds in histological images during histological image processing. A mask was first generated through edge and contour detection to separate tissue regions and backgrounds in histological images. In biomedical research, colored histological images are usually used to highlight the specific tissue morphology and cell structure. However, many edge detection methods have been developed for grayscale images, and RGB-to-gray conversion may cause information loss, including information on edges and borders. Therefore, in our method, the edges were detected directly from colored images through the large-color-difference formula (please see the Methods and Materials section for more details). Finally, the contours were detected from the image after the morphological operation. Although a small part of the structure in some results was lost (Figure 2a), subsequent processing steps were not seriously affected. Nearly complete tissue regions in histological images were segmented
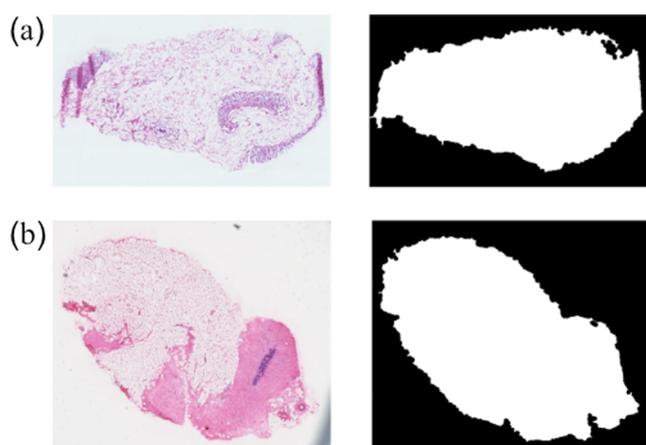
Figure 2. Successful histological image extractions with different tissue conditions, original histological images (on the left), and the corresponding mask images (on the right) after background removal. (a) One colorectal adenocarcinoma tissue sample in data set 3. (b) One breast cancer tissue sample in data set 2.
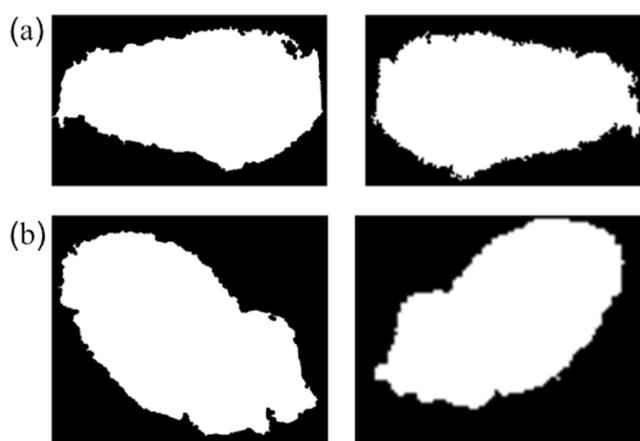

Figure 3. High similarity between tissue regions in histological mask images (on the left) and the results from on-tissue extraction of MSI data (on the right). (a) One colorectal adenocarcinoma tissue sample in data set 3. (b) One breast cancer tissue sample in data set 2.

through our method successfully. Even tissue containing many tiny empty spaces, such as adipose tissue, and most of the tissue borders were successfully segmented (Figure 2b).

The downsampling strategy reduced the computational cost and process time in edge and contour detection without significantly affecting subsequent processing. A high-resolution histological image often contains over one million pixels; thus, image processing requires many computational resources and a large amount of processing time. In the mask generation step, edge detection and contouring were performed on a downsampled histological image to reduce computational cost and accelerate processing. After a mask was generated, the mask was upsampled to the original size with nearest-neighbor interpolation. Although some background regions may be detected in tissue regions when using this strategy, the background regions do not affect subsequent processing. Compared to MSI data, histological images usually provide more precise border information for separating tissue regions and background, and background removal algorithms and methods in histology have been developed and applied for a while. Therefore, in the on-tissue spectra extraction step of MSI data, the contour from a histological image is used as a reference to determine which spectra belong to the tissue region.

On-Tissue Spectra Extraction. In our evaluation, the tissue region generated from MSI data on-tissue spectra extraction would be highly similar to the tissue region in the corresponding histological image (Figure 3) if there were no significant chemical noise-like carry-over signals in MSI data. The nontissue regions of MS images significantly affect registration accuracy and success and should be removed before image registration.[13] Previously, Race et al. proposed a method for removing nontissue regions using k-means clustering ($k = 3$) with the cosine distance.[16] Nevertheless, this method may not be suitable for all MSI data because of the variability of the MSI experiment. In this work, we optimized the method of Race et al. and utilized tissue contour information from histology as a reference to extract on-tissue spectra through shape matching. In shape-matching tasks, the use of Hu moments[17] is one method to evaluate the similarity between two shapes. Based on Hu invariants,[17] Hu moments

have been proven invariant to translation, scale, and rotation but variant to reflection. However, we found that reflection would not significantly affect the result of shape matching.

Hyperspectral Visualization. Most image registration algorithms are developed for one-to-one registration. However, many features with different intensity patterns exist in single MSI data. When representing the data, it is difficult to manually select a single ion image of a specific m/z value. Therefore, summarizing MSI data through multivariate techniques is necessary to acquire a single representative image. Many dimensionality reduction methods have been applied to MSI data, including nonlinear or linear methods.[18] Principal component analysis (PCA) is a common linear dimensionality reduction method that provides useful help in MSI data analysis. However, the linear relationship assumptions used in those methods may not be suitable for the inherent nonlinearity of the biological model. They may lose some underlying structure information in the MSI data.

Thus, many studies use nonlinear dimensionality reduction methods such as t-SNE and uniform manifold approximation and projection (UMAP) in biological analysis. Although all features of high-dimensional data can be embedded into two or three dimensions using t-SNE, providing more information in the visualization of MSI data and calculating the pairwise distance matrix between points requires a large computational source. It is unsuitable for data sets with large data points. The UMAP algorithm is scalable and suitable for hyperspectral data with many pixels and has been evaluated and applied in MSI.[19,20] Therefore, we evaluated two dimensionality reduction methods, PCA and UMAP. With our evaluation data, hyperspectral visualization images containing similar patterns with histology could be generated through both PCA and UMAP (Figure 4).

Multimodal Registration. In our evaluation, most representative MS images were successfully registered to correspond to histological images (Figure 5a,b). In MSIr, there are two steps in the image registration. First, the orientation and scale differences are processed, and second, accurate registration is performed. In the first step, the initial registration is used to determine the reflection and large-angle rotation differences and the spatial resolution relationship between the representative MS image and the histological image. After the first step, the representative MS and
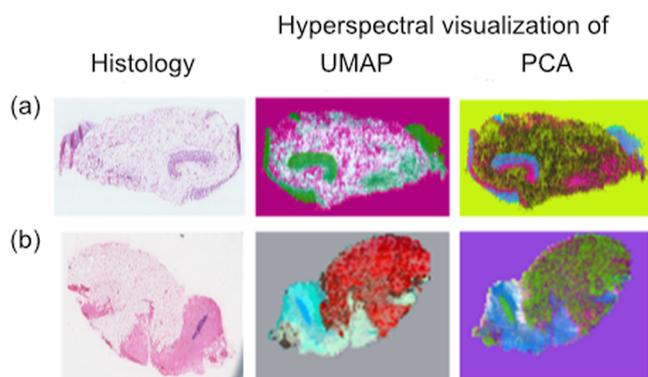
Figure 4. Highly similar patterns within histological images (on the left) and corresponding MS hyperspectral visualization images of UMAP (in the center) as well as PCA (on the right). (a) One colorectal adenocarcinoma tissue sample in data set 3. (b) One breast cancer tissue sample in data set 2.
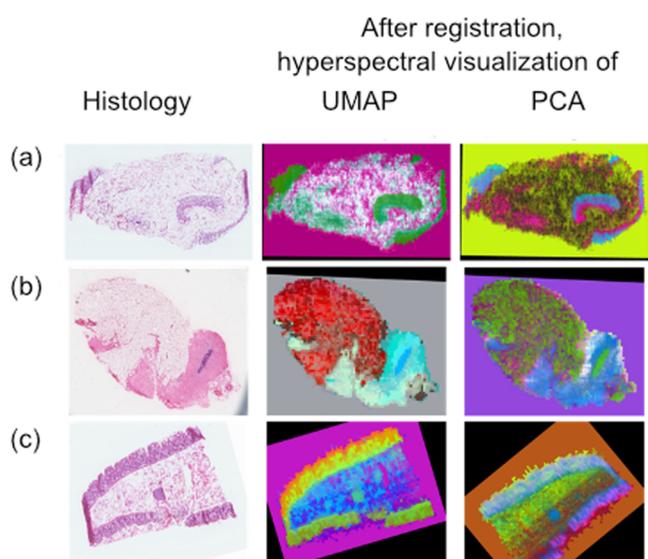

Figure 5. MS hyperspectral visualization images of UMAP (on the center) and PCA (on the right) after valid (a,b) and failed (c) registration and the corresponding histological images (on the left). (a) One colorectal adenocarcinoma tissue sample in data set 3. (b) One breast cancer tissue sample in data set 2. (c) Another colorectal adenocarcinoma tissue sample in data set 3.

histological images should have a close orientation and similar size. In the second step, intensity-based registration is implemented to accurately register the representative MS image to the histological image through Elastix after the initial registration. In our evaluation data, most of the data were successfully registered to correspond to histological images.

Nevertheless, several data had obviously failed registration, and the reasons for failed registration in those data occurred during the initial registration step. During the initial registration step, searches for possible orientation differences are performed using the mutual information loss between the MSI data and the histological image. The representative MS image could highly affect registration. For example, in one failed registration data set, MSI data were successfully registered to histology after dimensionality reduction using UMAP but unsuccessfully after dimensionality reduction using PCA (Figure 5c). Thus, in addition to the quality of MSI data, dimensionality reduction methods generating anatomical

structure information also influence the registration success. The capability of visualizing the nonlinear method (e.g., UMAP) can usually result in more successful registrations than other linear methods, including PCA, which is consistent with the observations from the work of Smets et al.[19]

In the evaluation results of MSIr, the mean Dice coefficients of the two-dimensionality reduction methods UMAP and PCA were 0.83 and 0.83, respectively, and the mean Hausdorff distances were 2.83 and 2.97 pixels, respectively. To evaluate the registration accuracy, the standard answer regions in each evaluation data set were generated based on the specific m/z value from the previous report of each study. The regions in histological images corresponding to standard answers in MSI data were also manually selected. The detailed effects using different dimensionality reduction methods are listed in Table 1.

Table 1. Effects Using Different Dimensionality Reduction Methods

dimensionality reduction method	UMAP	PCA
Dice Coefficient		
all data (n = 27)	[0.63, 0.97]	[0.54, 1.00]
exclude failed registrations (number of failed registrations)	[0.76, 0.90]	[0.76, 0.90]
Hausdorff Distance		
average in MSI pixels of all data (n = 27)	[0.47, 6.07]	[0.64, 6.58]
average in MSI pixels excluding failed registrations (number of failed registrations)	[1.49, 4.17]	[1.55, 4.39]

It was challenging to generate precise ground truth regions from MSI data because of the spatial resolution difference between MSI data and histology. Additionally, the signal in MSI data was probably suppressed or diffused, resulting from smaller or larger ground truth regions than corresponding regions in histological images. These would cause a lower Dice coefficient and a larger Hausdorff distance.

Recently, the application of machine learning and deep learning in microscopy has gained popularity, especially the application of classification and segmentation in histology, which is increasing significantly.[21] Through these techniques, the accuracies of different cancer type classifications are improved. However, due to the low chemical specificity of histology, classifications of some cancer subtypes based on only histology information using deep learning-based methods are challenging tasks.[22] In the published work of Race et al.,[16] the annotation transfer between MSI and histology showed that the histology to MSI annotation transfer assisted in the extraction of related ion images and statistical analysis of an region of interest (ROI) and that the MSI to histology annotation transfer helped the discovery of tumor heterogeneity using molecular information. MSI can provide information on the spatial distribution of molecules to complement the low chemical specificity of histology and make the performances of deep learning more powerful in cancer subtype classifications. Image registration would be necessary for data analysis between MSI and histology. The combination of MSI and histology should be more popular.

Web Interface. The home page of MSIr is shown in Figure 6. MSIr provides the services of automatic image registration between histology, MSI, and spectral index extraction. In automatic image registration, the user must upload centroided

Figure 6. Home page of MSIr.

MSI data in an imzML format file and corresponding histological image (png, jpg, and jpeg formats are available, brightness: 200 to 300, contrast: 40 to 70, see Figure S9) to MSIr and set the required parameters. UMAP and PCA are available for utilization in MSIr. Although the hyperspectral visualization images using PCA and UMAP both contain similar patterns to histology in our evaluation, we need to indicate that hyperspectral visualization images from multivariate techniques highly depend on the quality and status of MSI data. Users should select proper dimensionality reduction techniques for their MSI data. Masks can be drawn manually or automatically. In spectral index extraction, the user can select ROIs from a histological image on the ROI selection page (Figure 7). The spectral indices in the ROI based on registration results are extracted and output in text format.
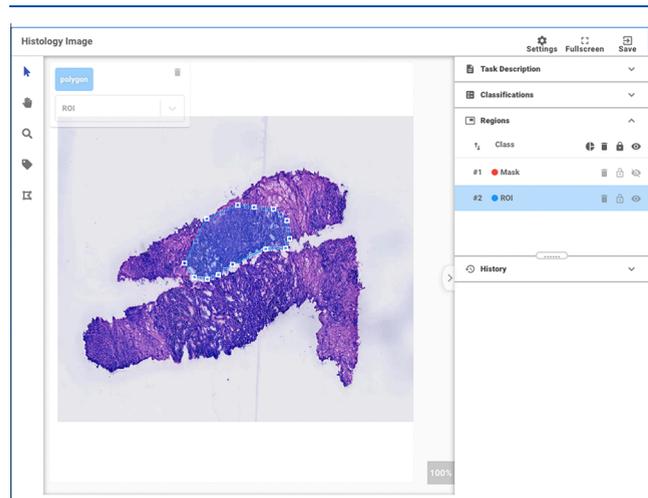

Figure 7. Mask drawing and ROI selection page.

Limitations. To perform registration successfully, the histological image should have a tissue border for initial image registration. Additionally, large and clear components should be able to be observed in both histological and MS images. Otherwise, the registration may fail due to insufficient information for initial registration and intensity-based registration. Although attempts are made to determine the orientation and scale differences between MSI data and histological images in the initial registration step by comparing the mutual information loss, this method will fail when the mutual information loss of eight orientation differences are close to each other. This occurs when the shape of the tissue is very close to being symmetrical. In the case of MSI results containing leaked analyst or image artifacts, image registration may be inaccurate due to the adjustment of the scale difference of the larger histological image with the leaked parts in the multimodal registration.

CONCLUSIONS

A web service, MSIr, was developed with functions of automatic image registration between MSI and histology and easy extraction of the corresponding spectrum from histological ROIs. Among 27 sample data from five studies, most registrations were effective without excessive shifts. This software can help users without image processing backgrounds or coding skills solve image registration problems in an easy-to-use manner. MSIr is freely available at https://msir.cmdm.tw, and the source code is available at https://github.com/CMDM-Lab/MSIr.

METHODS AND MATERIALS

Experimental Data Sets. Public data sets from four studies: (1) data set 1: colorectal adenocarcinoma study of Oetjen et al.,[23] (2) data set 2: breast cancer study of Guenther et al.,[24] (3) data set 3: colorectal adenocarcinoma study of Veselkov et al.,[12] and (4) data set 4: esophageal lymph node metastasis study of Abbassi-Ghadi et al.[25] Four data sets were collected from the MetaboLights database.[26] The public studies' experimental information about sample preparation and data acquisition can be found in the MetaboLights database. Data set 5[27] is the human breast cancer study provided by Prof. Cheng-Chih Hsu Research Group (Department of Chemistry, National Taiwan University). The information on these studies from the public database is listed in Table S1. All acquired data sets include MSI data and histological images.

Histological Image Processing. Histological image processing aims to remove the background and acquire tissue information for subsequent procedures. Edges in the histological image are detected through the algorithm based on the large-color-difference formula[28] to separate the tissue region and background. The downsampled histological image is used in edge detection to reduce the processing time. The resulting edge detection image is processed with a morphological operation (closing with a circle-shaped structural element) to remove tiny discontinuities between adjacent edges. After the morphological operation, the contours are detected from the resulting image, and the contour with the maximum area is taken as a tissue contour. If the area of any other contour is larger than half of the maximum contour area, the contour is also taken as a tissue contour. Based on tissue contours, a binary mask separating the tissue region and background is generated and upsampled to the original size of the histological image with nearest-neighbor interpolation. In the last step of histological image processing, the background in the histological image is processed to black based on a binary mask. All image processing procedures are implemented using the Python package "OpenCV-Python" (version 4.5.3.56).

MSI Data Processing. In the second procedure, mass spectrometry data processing is applied to summarize MSI data into a hyperspectral visualization image containing the anatomical information and contour information for the next procedure, multimodal registration. The imzML format file is parsed through the Python package "pyimzML" (version 1.5.1), and spectral data are organized into a 2D pixel-m/z format in a sparse matrix. Data are normalized by total ion

count normalization. The on-tissue spectra extraction described below is applied to remove the background region spectral effect. Peak picking using the Python package "ms_peak_picker" (version 0.1.33) is performed on the mean spectrum of on-tissue spectra to determine the locations and intensities of peaks. Peak intensities are generated by summing the intensity values of adjacent peaks whose m/z difference from the m/z value of the main peak is lower than the threshold value. In the last step of MSI data processing, MSI data are subjected to nonlinear dimensionality reduction for hyperspectral visualization.

On-tissue spectra extraction is implemented to determine the spectra in the tissue region. The method of on-tissue spectral extraction is modified from the method proposed by Race et al.[16] The spectra of all pixels are clustered by a k-means algorithm with the cosine distance metric, and clustering of different k values from two to four operate sequentially until acquiring the appropriate tissue contour. The clustering result of each round of clustering is reshaped to the 2D image form first. The clustering labels of pixels in the image border are counted, and the label with the maximum counts is seen as belonging to the background region. A combination set of those labels will be generated if a clustering label is not classified as background. For this combination set, pixels with any labels from the combination were temporarily taken as tissue regions. The contours of this temporary tissue region are detected. The dissimilarity between the detected contours of histological images and tissue contours is calculated based on Hu moments.[17] Suppose the dissimilarity is lower than the threshold (threshold = $0.13 - (k - 1) \times 0.005$, k: the k value of the k-means algorithm). In this case, the spectra of pixels in this temporary tissue region are taken as on-tissue spectra, and the on-tissue spectra extraction is complete. If the dissimilarities of all combinations from the combination set are higher than the threshold, the next round of clustering will be performed. If the dissimilarities of all combinations in every clustering round are higher than the threshold, the result with the smallest dissimilarity is used.

Hyperspectral visualization is used to summarize complex MSI data to generate a summary image with anatomical information. MSI data, including on-tissue and background spectra, are embedded into three dimensions through dimensionality reduction methods. The background spectrum is represented by the mean spectrum of the spectra in the background region. The embedding results of dimensionality reduction are finally converted to the RGB color space. The selectable dimensionality reduction methods include UMAP[29] and PCA. UMAP is implemented using the Python package "umap-learn" (version 0.5.1), and PCA is applied using Python package "scikit-learn" (version 0.24.2).

Multimodal Registration. The third procedure, multimodal registration, is implemented to acquire the optimal transformation matrices for registering MSI data into the corresponding histological image. The MS hyperspectral visualization image and the processed histological image are first converted from a 3D RGB image into a 1D grayscale image using RGB-to-gray conversion.

In the initial image registration step, the orientation and scale differences between the histological and MSI hyperspectral visualization images are processed. Eight orientation difference situations are generated and evaluated to determine the orientation differences. The eight situations are large-angle rotation differences (0, 90, 180, and 270°) with and without

reflection. Mutual information loss between two modality images is used as a matching image metric to determine the possible difference. To solve the spatial resolution difference between the two images, the scale factor is calculated based on the circumcircle radius of the tissue contours in the two images. The temporary histological image is scaled with the reciprocal of the scale factor. The eight situations are separately generated through rotation and mirror flip operations. To minimize the deviation from relative displacement under any situation, the centers of mass of tissue regions in two images are overlapped through translation. The image matching metrics of all situations are compared, and the situation with minimal mutual information loss is taken as the possible orientation difference between the two images. The MS hyperspectral visualization image is transformed with mirror flip, rotation, and scaling operations based on a possible orientation difference.

The small difference between the hyperspectral visualization image after initial image registration and the histological image is eliminated in intensity-based image registration. Intensity-based image registration is performed using the Python package "itk-elastix" (version 0.13.0), an ITK Python interface to Elastix.[30] Grayscale and corresponding mask images from MSI data and histology are input. An adaptive stochastic gradient descent optimizer is used to optimize the optimal transform parameters between moving and fixed images to minimize the mutual information loss function. Histological images are used as the fixed images, and MS images are the moving images.

In the last step of multimodal registration, a series of transformation parameters are integrated into three transformation matrices, including the matrix for processing the situations of reflection relationship and large-angle rotation difference, the scale matrix, and the similarity transformation matrix converted from transformation parameters in Elastix. The three transformation matrices are output in a single text format file.

Spectral Index Extraction. In the spectral index extraction, MSI data spectral indices in the ROI of the corresponding histological image are extracted and output to users for subsequent analysis. ROIs are used to extract the corresponding spectral indices in MSI data based on the transformation matrices from automatic registration. To determine the corresponding ROI spectra, the index table in 2D form is transformed sequentially with nearest-neighbor interpolation by three transformation matrices from multimodal registration. To verify whether the spectrum was in the ROI, the area of each spectrum at the ROI and the total area of each spectrum at the histological image scale are calculated. If the ROI and total area percentage at the histological image scale are larger than the threshold (50%), the spectrum is identified as one of the ROI spectra. Finally, the spectrum indices at the ROI are output in a text format.

Registration Evaluation. A total of 27 sample data points are selected from five data sets to evaluate the performance of MSIr. The sample data for evaluation are selected based on obvious visual inspection and feasible separation from the neighborhood in both the MSI data and histological images. The region as a standard answer in MSI data is generated based on intensity data of specific m/z values. For each specific m/z value, the intensity data are rescaled into a range of 0−255 with a low quantile threshold of 0% and a high quantile threshold of 99%. Then, each m/z value ion image is merged

with the average weight, and this merged image is binarized through the Otsu threshold. Depending on the status of binary images, morphological operations (closing with a 3×3 square structural element) are applied to remove small holes. The specific m/z values of each study are listed in Table S2, and the specific m/z values are selected based on significant findings of metabolites in each of the previous publications (data sets 1–3) or manually selected based on the features from histological images for data set 4 and data set 5. The regions in histological images corresponding to the standard answer region in MSI data are labeled manually. These regions include intact tissue regions, some tissue regions with abnormal tissue types, and holes in tissue regions.

In the evaluation of registration results, the regions in the registration result corresponding to the labeled regions in the histological image are extracted. Both the Dice coefficient and Hausdorff distance are used to evaluate the MSI space because these two coefficients have been used for registration evaluation between MSI data and histology.[13,16]

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.2c04360.

> Overview of experimental data set information, m/z values used in each study, detailed workflow of MSIr, home page and data set page of MSIr, ROI selection page, and brightness and contrast distribution of the histology images of successful registration (PDF)

AUTHOR INFORMATION

Corresponding Author

Yufeng Jane Tseng − *Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei 10617, Taiwan; The Metabolomics Core Laboratory, Centers of Genomic and Precision Medicine and Department of Computer Science and Information Engineering, National Taiwan University, Taipei 10617, Taiwan; School of Pharmacy, College of Medicine, National Taiwan University, Taipei 10002, Taiwan;* ⓘ orcid.org/0000-0002-8461-6181; Phone: +886.2.3366.4888#529; Email: yjtseng@csie.ntu.edu.tw; Fax: +886.2.23628167

Authors

Bo-Jhang Lin − *Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei 10617, Taiwan*

Tien-Chueh Kuo − *The Metabolomics Core Laboratory, Centers of Genomic and Precision Medicine, National Taiwan University, Taipei 10617, Taiwan;* ⓘ orcid.org/0000-0002-4794-0134

Hsin-Hsiang Chung − *Department of Chemistry, National Taiwan University, Taipei 10617, Taiwan;* ⓘ orcid.org/0000-0001-6476-3814

Ying-Chen Huang − *Department of Chemistry, National Taiwan University, Taipei 10617, Taiwan*

Ming-Yang Wang − *Department of Surgery, National Taiwan University Hospital, Taipei 10002, Taiwan*

Cheng-Chih Hsu − *Department of Chemistry, National Taiwan University, Taipei 10617, Taiwan;* ⓘ orcid.org/0000-0002-2892-5326

Po-Yang Yao − *Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei 10617, Taiwan*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.analchem.2c04360

Author Contributions

Y.J.T. conceived the project. B.J.L. and T.C.K. designed the algorithm. H.H.C., Y.C.H., and C.C.H. performed the MSI experiments. M.Y.W. collected and provided tissue samples. B.J.L., T.C.K., and P.Y.Y. developed the website. All authors wrote the manuscript. All authors have reviewed and approved the final version of the manuscript.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Buchberger, A. R.; DeLaney, K.; Johnson, J.; Li, L. *Anal. Chem.* **2018**, *90*, 240.

(2) Ryan, D. J.; Spraggins, J. M.; Caprioli, R. M. *Curr. Opin. Chem. Biol.* **2019**, *48*, 64−72.

(3) Swales, J. G.; Hamm, G.; Clench, M. R.; Goodwin, R. J. A. *Int. J. Mass Spectrom.* **2019**, *437*, 99−112.

(4) Perez, C. J.; Bagga, A. K.; Prova, S. S.; Yousefi Taemeh, M. Y.; Ifa, D. R. *Rapid Commun. Mass Spectrom.* **2019**, *33*, 27−53.

(5) Brown, H. M.; McDaniel, T. J.; Fedick, P. W.; Mulligan, C. C. *Anal. Methods* **2020**, *12*, 3974−3997.

(6) Parrot, D.; Papazian, S.; Foil, D.; Tasdemir, D. *Planta Med.* **2018**, *84*, 584.

(7) Tuck, M.; Blanc, L.; Touti, R.; Patterson, N. H.; Van Nuffel, S.; Villette, S.; Taveau, J. C.; Römpp, A.; Brunelle, A.; Lecomte, S.; Desbenoit, N. *Anal. Chem.* **2021**, *93*, 445−477.

(8) Porta Siegel, T.; Hamm, G.; Bunch, J.; Cappell, J.; Fletcher, J. S.; Schwamborn, K. *Mol. Imaging Biol.* **2018**, *20*, 888−901.

(9) Iakab, S. A.; Ràfols, P.; Correig-Blanchar, X.; García-Altares, M. *Anal. Chem.* **2021**, *93*, 6301−6310.

(10) Arentz, G.; Mittal, P.; Zhang, C.; Ho, Y.-Y.; Briggs, M.; Winderbaum, L.; Hoffmann, M.; Hoffmann, P. *Adv. Cancer Res.* **2017**, *134*, 27−66.

(11) Patterson, N. H.; Tuck, M.; Van de Plas, R.; Caprioli, R. M. *Anal. Chem.* **2018**, *90*, 12395−12403.

(12) Veselkov, K. A.; Mirnezami, R.; Strittmatter, N.; Goldin, R. D.; Kinross, J.; Speller, A. V.; Abramov, T.; Jones, E. A.; Darzi, A.; Holmes, E.; Nicholson, J. K.; Takats, Z. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 1216−1221.

(13) Abdelmoula, W. M.; Škrášková, K.; Balluff, B.; Carreira, R. J.; Tolner, E. A.; Lelieveldt, B. P. F.; van der Maaten, L.; Morreau, H.;

van den Maagdenberg, A. M. J. M.; Heeren, R. M. A.; McDonnell, L. A.; Dijkstra, J. *Anal. Chem.* **2014**, *86*, 9204−9211.

(14) Patterson, N. H.; Yang, E.; Kranjec, E. A.; Chaurand, P. *Bioinformatics* **2019**, *35*, 1261−1262.

(15) Bemis, K. D.; Harry, A.; Eberlin, L. S.; Ferreira, C.; van de Ven, S. M.; Mallick, P.; Stolowitz, M.; Vitek, O. *Bioinformatics* **2015**, *31*, 2418−2420.

(16) Race, A. M.; Sutton, D.; Hamm, G.; Maglennon, G.; Morton, J. P.; Strittmatter, N.; Campbell, A.; Sansom, O. J.; Wang, Y.; Barry, S. T.; Takáts, Z.; Goodwin, R. J. A.; Bunch, J. *Anal. Chem.* **2021**, *93*, 3061−3071.

(17) Hu, M. K. *IRE Trans. Inf. Theory* **1962**, *8*, 179−187.

(18) Verbeeck, N.; Caprioli, R. M.; Van de Plas, R. *Mass Spectrom. Rev.* **2020**, *39*, 245−291.

(19) Smets, T.; Verbeeck, N.; Claesen, M.; Asperger, A.; Griffioen, G.; Tousseyn, T.; Waelput, W.; Waelkens, E.; De Moor, B. *Anal. Chem.* **2019**, *91*, 5706−5714.

(20) Smets, T.; Waelkens, E.; De Moor, B. *Anal. Chem.* **2020**, *92*, 5240−5248.

(21) Kobayashi, S.; Saltz, J. H.; Yang, V. W. *World J. Gastroenterol.* **2021**, *27*, 2545−2575.

(22) Tran, K. A.; Kondrashova, O.; Bradley, A.; Williams, E. D.; Pearson, J. V.; Waddell, N. *Genome Med.* **2021**, *13*, 152.

(23) Oetjen, J.; et al. *Gigascience* **2015**, *4*, 20.

(24) Guenther, S.; Muirhead, L. J.; Speller, A. V. M.; Golf, O.; Strittmatter, N.; Ramakrishnan, R.; Goldin, R. D.; Jones, E.; Veselkov, K.; Nicholson, J.; Darzi, A.; Takats, Z. *Cancer Res.* **2015**, *75*, 1828−1837.

(25) Abbassi-Ghadi, N.; Veselkov, K.; Kumar, S.; Huang, J.; Strittmatter, N.; Guenther, S.; Kudo, H.; Goldin, R.; Takats, Z.; Hanna, G. B. *Chem. Commun.* **2014**, *50*, 3661.

(26) Haug, K.; Cochrane, K.; Nainala, V. C.; Williams, M.; Chang, J. K.; Jayaseelan, K. V.; O'Donovan, C. *Nucleic Acids Res.* **2020**, *48*, D440−D444.

(27) Lin, L.-E.; Chen, C.-L.; Huang, Y.-C.; Chung, H.-H.; Lin, C.-W.; Chen, K.-C.; Peng, Y.-J.; Ding, S.-T.; Wang, M.-Y.; Shen, T.-L.; Hsu, C.-C. *Anal. Chim. Acta* **2020**, *1100*, 75−87.

(28) Abasi, S.; A. Tehran, M. A.; Fairchild, M. D. *Color Res. Appl.* **2020**, *45*, 632−643.

(29) McInnes, L.; Healy, J.; Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. **2018**, arXiv:1802.03426. arXiv preprint.

(30) Klein, S.; Staring, M.; Murphy, K.; Viergever, M. A.; Pluim, J. P. W. *IEEE Trans. Med. Imag.* **2010**, *29*, 196−205.