Sequential paired covariance for improved visualization of mass spectrometry imaging datasets

Crystal L. Pace | Kenneth P. Garrard | David C. Muddiman

Abstract

Untargeted analyses in mass spectrometry imaging produce hundreds of ion images representing spatial distributions of biomolecules in biological tissues. Due to the large diversity of ions detected in untargeted analyses, normalization standards are often difficult to implement to account for pixel-to-pixel variability in imaging studies. Many normalization strategies exist to account for this variability, but they largely do not improve image quality. In this study, we present a new approach for improving image quality and visualization of tissue features by application of sequential paired covariance (SPC). This approach was demonstrated using previously published tissue datasets such as rat brain and human prostate with different biomolecules like metabolites and N-linked glycans. Data transformation by SPC improved ion images resulting in increased smoothing of biological features compared with commonly used normalization approaches.

Keywords

data visualization, heatmaps, mass spectrometry imaging, MSIReader, sequential paired covariance

1 | INTRODUCTION

Most mass spectrometry imaging (MSI) methodologies face pixel-to-pixel or voxel-to-voxel variability challenges due to changes in tissue density across a biospecimen and heterogeneity of matrix application. This can result in pixelated ion images making it difficult to delineate edges of histological features. To account for this, molecular standards with similar ionization efficiencies to the analyte of interest can be sprayed underneath biological samples and used for normalization. However, normalization standards are difficult to choose for untargeted experiments because of the variability of ionization efficiencies between the molecules detected. Many post-processing normalization methods available do not require a standard to reduce the relative standard deviation between pixels/voxels. These include normalization to the total ion current (TIC), median, or to the sum of biological peaks. However, these normalization techniques do not always reduce the pixelated appearance of ion images, and thus, the need for additional approaches still exists.

Sequential paired covariance (SPC) was introduced for mass spectrometry applications in 1995 to improve signal-to-noise ratios in liquid separations coupled with mass spectrometry datasets. This ultimately allowed for rapid assessment of significant peaks in the electropherogram. The SPC algorithm represents the dot product of sequential mass spectra, thus reducing the effect of variable noise peaks in the dataset. Here, we present a new data transformation specifically to improve the visualization of ion images in MSI methodologies in order to easily decipher between histological features of a tissue. To our knowledge, this is the first application of SPC to MSI datasets for improving ion images. We demonstrate this improvement for a variety of heterogeneous datasets which...
were chosen as they each have distinct biological features showing that SPC allows for easy interpretation and visualization of colocalized features. This feature has been incorporated into MSiReader so that it can be easily applied to any MSI dataset and will be available in the next release of MSiReader for the MSI community.11,12

2 | METHODS AND MATERIALS

2.1 | Calculation of SPC in MSiReader

Data transformation with SPC was implemented in MSiReader to calculate across large imaging datasets. Before calculation of SPC, voxels containing zeros were replaced with not-a-number (NaN) to omit them from SPC calculations. This step was necessary to prevent the propagation of zero values throughout the ion image. SPC was then calculated by taking a mathematical algorithm of adjacent and diagonal voxels (Figure 1). Four algorithmic variations were implemented into MSiReader and compared in this manuscript: product (pSPC), sum (sumSPC), median (medSPC), and midpoint (midSPC). For all voxels except those touching the border of the image or region of interest, the SPC calculation results in the abundance of 9 voxels being transformed. Depending on the absolute abundance of the ion, this results in a significant magnification of ion abundances for the pSPC function with a large dynamic range between the lowest and highest value. Therefore, it was necessary to also include the option for logarithmic transformations to visualize the distribution of ions throughout the tissue samples.

2.2 | Evaluation of imaging datasets

A few datasets were selected to evaluate the use of SPC to improve the visualization of spatial distributions. These datasets were chosen based on the heterogeneous nature of the biological tissues. Each dataset was obtained by IR-MALDESI MSI coupled to an Orbitrap mass spectrometer as previously described.13,14 Thermo .RAW data files were converted to mzML files using MSConvert15 and then to imzML files using imzMLconverter.16 All ion images were generated and analyzed using MSiReader, a free MATLAB software for imaging analyses.11,12

2.2.1 | MSI of metabolites in rat brain

Rat brain sectioned to a thickness of 12 μm was placed inside the sample enclosure purged with nitrogen gas (Arc3 Gases, Raleigh, NC, USA), and a thin ice layer was formed by controlled exposure to humidity. An electro spray ionization source was achieved with 50% optima LC–MS grade methanol (Fisher Scientific Hampton, NH, USA) and 0.2% formic acid (Sigma-Aldrich, St. Louis, MO, USA) solution at a flow rate of 1.5 μl/min and spray voltage of 4000 V. A mid-IR laser Opolette 2731/3034 (Opotek, Carlsbad, CA, USA) operating at 2.94 μm was used for laser ablation at 1.1 mJ per voxel with a spatial resolution of 150 μm. IR-MALDESI was coupled to a Q Exactive Plus mass spectrometer (Thermo Fisher Scientific, Breman, Germany) for high-resolution accurate mass spectrometry at a resolution of 140,000FWHM at 200 m/z. Automatic gain control (AGC) was disabled and a fixed injection time of 25 ms was set to synchronize the laser ablation event with collection of ions in the C-trap.

2.2.2 | MSI of N-linked glycans in human prostate

FFPE human prostate tissue was dewaxed/delipidated, antigen retrieved, and enzymatically cleaved by PNGase F for the analysis of N-linked glycans. The digested human prostate tissue was placed inside the sample enclosure, and a thin ice layer was formed by controlled exposure to humidity. An electro spray ionization source was achieved with 60% acetonitrile and 1 mM acetic acid solution at a flow rate of 2 μl/min and spray voltage of 3600 V for negative ionization mode. A mid-IR laser (JGM Associates, Burlington, MA, USA) operating at 2.97 μm was used for laser ablation at 1.2 mJ per voxel with a spatial resolution of 150 μm. IR-MALDESI was coupled to an Orbitrap Exploris 240 mass spectrometer (Thermo Fisher Scientific, Breman, Germany) for high-resolution accurate mass spectrometry at a resolving power of 240,000FWHM at 200 m/z. AGC was disabled and a fixed injection time of 90 ms was set to synchronize the laser ablation event with collection of ionized glycans in the C-trap.

3 | RESULTS AND DISCUSSION

Before demonstrating data transformation with SPC on real biological data where absolute concentrations on biomolecules in all tissue

![FIGURE 1](image-url) A simple 4 × 4 region of interest showing the calculation of pSPC in MSiReader
features is unknown, SPC was demonstrated on theoretical data to show improved image quality. Figure 2A represents theoretical features where the abundance of specific biomolecules should be consistent across that particular feature. However, in a typical MSI approach, we would expect to observe pixel-to-pixel variability as seen in Figure 2B. These values were calculated using the rand() function in Excel. These images largely correspond to experimental data observed by MSI approaches. Transforming the data with pSPC

![Theoretical Image](image1)
![Expected Pixel Variability](image2)
![Ln(pSPC(Pixel Variability))](image3)

**FIGURE 2** Demonstration of improved image quality by pSPC using theoretical data

![Ion images of metabolites in the rat brain](image4)

**FIGURE 3** Ion images of metabolites in the rat brain (A) before and (B) after pSPC. The transformation with pSPC visually smooths out the pixelated nature and allows for easier visualization of histological features such as the corpus callosum. (C) Zoomed-in features of the corpus callosum are shown to observe a higher level of detail when transformed with the pSPC algorithm.
shown in Figure 2C demonstrates the ability to clean up the pixelated nature of ion images by smoothing the data into cleaner features.

3.1 | MSI of metabolites in rat brain

To evaluate the use of SPC on heterogeneous tissue, we investigated a biological tissue from a previous study where direct comparisons between biological features were made. Four distinct features of the rat brain, such as the cerebral cortex, corpus callosum, caudate putamen, and septal region were compared for differences in metabolite profiles and abundances. From this study, we found significant differences particularly between the corpus callosum and other analyzed regions (Figure 3A). From these ion images, the global differences throughout the tissue features can be visualized. However, due to the pixelated nature of the images, the edges of many of the histological features are not smooth, and this makes it difficult to decipher the biological features. When the pSPC algorithm is applied to each ion image, the noisy pixelated features are reduced. The distinct histological features and ion distributions can then be visualized much more clearly (Figure 3B).

It is important that the data are not altered such that different biological conclusions are made in the process of applying the SPC algorithm. To validate that relative relationships remained the same, we exported the abundances in each of the histological regions and calculated an average abundance per region for both the raw and the pSPC-transformed datasets. Figure 4A show bar graphs of the average abundance of cholesterol in each brain region. While the magnitude between raw abundances and SPC values are drastically different, the relative distributions of abundances among the regions of interest remained the same. Region 4 (the corpus callosum) had the highest abundance of cholesterol while Region 2 (cerebral cortex) remained the region with the lowest abundance of cholesterol. This confirmed the spatial relationships between these regions remained accurate after postprocessing with the pSPC algorithm. Relative standard deviation was calculated to estimate the variance independent of magnitude between the natural log of raw abundances and pSPC transformations. Figure 4B represents the %RSD for cholesterol in each of the histological features of the brain showing a significant decrease in the variability when using pSPC.

3.2 | MSI of N-linked glycans in human prostate

We also applied SPC to a large dataset of N-linked glycans in the human prostate tissue (Figure 5A). This dataset includes glycans localized within small morphological features demonstrating the need for data transformation to visualize those specific features in the prostate tissue. The glycan in Figure 5B (757.2151 m/z), for example, is distributed across the majority of the tissue but absent in many of the small glandular features. Due to the pixelated nature of the ion image, many of those glandular features are hard to distinguish. The large circular stromatic feature in the top right corner is more easily observed to overlap with the ion image after transformation with pSPC. The glycan in Figure 5C (820.2117 m/z) was observed in the large glandular features in the top right while Figure 5D (1037.8650 m/z) represents a glycan particularly distributed throughout the prostate tissue with a decrease in abundance in the large circular feature in the top right. The black stripes seen across the ion images represent a loss of electrospray ionization during collection of data resulting in low abundant and zero-valued voxels. While these are present in the ion images before transformation by pSPC, they appear more prominent in the SPC images. This is a result of fewer voxels being multiplied (zero voxels receive a NaN to prevent zero propagation), thus blending the transition between zero-valued voxels and higher abundant values by 1 voxel in each direction.
3.3 | Comparison to other techniques

The gold standard for improving image resolution and accounting for voxel-to-voxel variability in MSI methodologies is normalization to a stable isotope labeled standard with similar ionization efficiency to the molecule of interest. For untargeted analyses, this becomes quite challenging as there are hundreds of metabolites detected in a single study and researchers may not know a priori which molecules will be of interest until after the untargeted analysis. Without internal standards, normalization strategies are heavily relied on to account for voxel variability. Common normalization strategies are demonstrated in Figure 6. Normalization to the TIC is the most common strategy where the abundance of an ion is divided by the sum of all ions in each individual voxel. Similarly, normalization to the midpoint, mean, max, and local TIC are also available options. Comparing the images of 757.2151 m/z in Figure 6 with the pSPC-transformed image in Figure 5B, one can see a significant improvement in image quality by application of pSPC.

Throughout this manuscript, we demonstrated the use of pSPC by calculating a product of abundances between adjacent and diagonal voxels. Additional approaches to calculating sequential covariance also include calculation of the sum, median, or midpoint between adjacent and diagonal voxels. Median image filtering is one of the most common ways to remove “salt and pepper” pixels from digital images. Examples of improvement with these additional approaches are shown in Figure 7 with 757.2151 m/z in the prostate tissue. Similar results were observed between product, sum, median, and midpoint with minor differences between all approaches. Where the pSPC approach suffered from a large magnitude of values requiring a logarithmic transformation, the sumSPC, medSPC, and midSPC approaches avoided this issue altogether. The midSPC approach appeared to create large voxel artifacts across the image while the other SPC approaches resulted in better smoothing across the features of the prostate tissue. Where the pSPC results in large, oversaturated tissue features and the medSPC results in undersaturated tissue features, the sumSPC approach results in improved image smoothing without under or oversaturation of tissue features. All four SPC algorithms with an optional logarithmic transformation are coded in MSiReader.

The decision on which SPC tool to use will likely be on a case-by-case basis for tissue type as well as complexity of individual spatial distributions. Because the prostate tissue has smaller features, the midSPC tool is less ideal as it creates large artifacts within the dataset. For a dataset with larger tissue features, this issue may not
interfere with the interpretation of the spatial distributions and could still be a useful approach for creating smoother ion images. Additionally, if a particular ion type has a large dynamic range across the tissue, the pSPC tool may be less ideal as using the logarithmic function and the min/max intensity bars to optimize the best scale could take some additional time. This was not an issue with the prostate tissue as most of the N-linked glycans were already low abundant signals. However, cholesterol in the rat brain did have a large dynamic range requiring the logarithmic function as well as scale bar optimization.
4 | CONCLUSIONS

Application of SPC was used to distinguish biological features in ion images more clearly. The SPC was applied to datasets with varying types of heterogeneous features. Analysis of average abundances before and after transformation showed that relative relationships between biological features were preserved despite application of the algorithm. The product SPC approach was compared with common normalization strategies for untargeted analyses showing that the pSPC transformation largely improves image quality over normalization techniques. This tool implemented into MSiReader allows for easy application and visualization of ion images.

ACKNOWLEDGEMENTS

IR-MALDESI-MSI measurements were carried out in the Molecular Education, Technology, and Research Innovation Center (METRIC) at North Carolina State University, which is supported by the State of North Carolina. The authors gratefully acknowledge the financial support received from the National Institutes of Health (R01GM087964 and R56AG063885) and North Carolina State University.

DATA AVAILABILITY STATEMENT

Metabolomic data is publicly available in METASPACE.

ORCID

Crystal L. Pace https://orcid.org/0000-0002-9339-6467
Kenneth P. Garrard https://orcid.org/0000-0002-0654-4776
David C. Muddiman https://orcid.org/0000-0003-2216-499X

REFERENCES

1. Tobias F, Hummon AB. Considerations for MALDI-based quantitative mass spectrometry imaging studies. J Proteome Res. 2020;19(9):3620-3630. doi:10.1021/acs.jproteome.0c00443
2. Deininger SO, Cornett DS, Paape R, et al. Normalization in MALDI-TOF imaging datasets of proteins: practical considerations. Anal Bioanal Chem. 2011;401(1):167-181. doi:10.1007/s00216-011-0429-z
3. Mellingr AL, Garrard KP, Khodjaniyazova S, Rabbani ZN, Gamcsik MP, Muddiman DC. Multiple infusion start time mass spectrometry imaging of dynamic SIL-glutathione biosynthesis using infrared matrix-assisted laser desorption electrospray ionization. J Proteome Res. 2022;21(3):747-757. doi:10.1021/acs.jproteome.1c00636
4. Källback P, Shariatgorji M, Nilsson A, Andrén PE. Novel mass spectrometry imaging software assisting labeled normalization and quantification of drugs and neuropeptides directly in tissue sections. J Proteomics. 2012;75(16):4941-4951. doi:10.1016/j.jprot.2012.07.034
5. Prentice BM, Chumbley CW, Caprioli RM. Absolute quantification of rifampicin by MALDI imaging mass spectrometry using multiple TOF/TOF events in a single laser shot. J Am Soc Mass Spectrom. 2017;28(1):136-144. doi:10.1007/s13361-016-1501-2
6. Chumbley CW, Reyzer ML, Allen JL, et al. Absolute quantitative MALDI imaging mass spectrometry: a case of rifampicin in liver tissues. Anal Chem. 2016;88(4):2392-2398. doi:10.1021/acs.analchem.5b04409
7. Tu A, Muddiman DC. Systematic evaluation of repeatability of IR-MALDESI-MS and normalization strategies for correcting the analytical variation and improving image quality. Anal Bioanal Chem. 2019;411(22):5729-5743. doi:10.1007/s00216-019-01953-5
8. Fonville JM, Carter C, Cloarec O, et al. Robust data processing and normalization strategy for MALDI mass spectrometric imaging. Anal Chem. 2012;84(3):1310-1319. doi:10.1021/ac201767g
9. Muddiman DC, Rockwood AL, Gao Q, et al. Application of sequential paired covariance to capillary electrophoresis electrospray ionization time-of-flight mass spectrometry: unraveling the signal from the noise in the electropherogram. Anal Chem. 1995;67(23):4371-4375. doi:10.1021/ac00119a027
10. Muddiman DC, Huang BM, Anderson GA, et al. Application of sequential paired covariance to liquid chromatography-mass spectrometry data. Enhancements in both the signal-to-noise ratio and the resolution of analyte peaks in the chromatogram. J Chromatogr A. 1997;771(1-2):1-7. doi:10.1016/S0021-9673(97)00069-1
11. Robichaud G, Garrard KP, Barry JA, Muddiman DC. MSiReader: an open-source interface to view and analyze high resolving power MS imaging files on Matlab platform. J Am Soc Mass Spectrom. 2013;24(5):718-721. doi:10.1074/jasms.13-01406-2
12. Bokhart MT, Nazari M, Garrard KP, Muddiman DC. MSiReader v1.0: evolving open-source mass spectrometry imaging software for targeted and untargeted analyses. J Am Soc Mass Spectrom. 2018;29(1):8–16. doi:10.1007/s13361-017-1809-6
13. Pace CL, Simmons J, Kelly RT, Muddiman DC. Multimodal mass spectrometry imaging of rat brain using IR-MALDESI and NanoPOTS-LCMS/MS. J Proteome Res. 2022;21(3):713-720. doi:10.1021/acs.jproteome.1c00641
14. Pace CL, Angel PM, Drake RR, Muddiman DC. Mass spectrometry imaging of N-linked glycans in a formalin-fixed paraffin-embedded human prostate by infrared matrix-assisted laser desorption electrospray ionization. J Proteome Res. 2022;21(1):243-249. doi:10.1021/acs.jproteome.1c00822
15. Kessner D, Chambers M, Burke R, Agus D, Mallick P. ProteoWizard: open source software for rapid proteomics tools development. Bioinformatics. 2008;24(21):2534-2536. doi:10.1093/bioinformatics/btn323
16. Race AM, Styles IB, Bunch J. Inclusive sharing of mass spectrometry imaging data requires a converter for all. J Proteomics. 2012;75(16):5111-5112. doi:10.1016/j.jprot.2012.05.035
17. Goyal A, Bijalwan A, Chowdhury M. A comprehensive review of image smoothing techniques. Int J Adv Res Comput Eng Technol. 2012;1(4):315-319.

How to cite this article: Pace CL, Garrard KP, Muddiman DC. Sequential paired covariance for improved visualization of mass spectrometry imaging datasets. J Mass Spectrom. 2022; 57(7):e4872. doi:10.1002/jms.4872