

Quimiometria, imatges hiperespectrals i teixit vegetal: fem ús de tota la informació de les nostres dades?

Chemometrics, hyperspectral images and plant tissue: Do we use all the information in our data?

Adrián Gómez-Sánchez,¹ Mónica Marro,² María Marsal,² Pablo Loza-Álvarez² i Anna de Juan¹

¹ Universitat de Barcelona. Grup de Quimiometria

² Barcelona Institute of Science and Technology (BIST). Institut de Ciències Fotòniques (ICFO)

Resum: Descobrir quins compostos químics hi ha en un teixit biològic i on estan localitzats ens dona una informació crucial sobre la funcionalitat d'aquests compostos i del mateix teixit. Aquesta informació la proporcionen les tècniques d'imatgeria hiperespectral (HI), que associen la informació espectral a la posició dels píxels a la mostra. Aquestes tècniques, però, generen una gran quantitat de dades complexes que s'han de gestionar amb l'ús d'eines quimiomètriques, com ara el mètode de resolució multivariant de corbes per mínims quadrats alternats (MCR-ALS), per poder fer una interpretació correcta de la mesura. A més, la informació complementària proporcionada per les diferents plataformes espectroscòpiques d'imatgeria hiperespectral es pot combinar mitjançant estratègies de fusió d'imatges per ajudar a obtenir una descripció més completa del teixit en estudi. Aquest treball s'orienta a mostrar el potencial de la combinació de la imatgeria hiperespectral i la quimiometria per interpretar la informació associada als teixits biològics. Per il·lustrar-ho, es mostra l'estudi de diverses microseccions de fulles de planta d'arròs mitjançant les imatges Raman, de fluorescència i la seva fusió.

Paraules clau: Imatgeria hiperespectral, fusió d'imatges, quimiometria, teixit vegetal.

Abstract: Discovering the nature and location of chemical compounds in a biological tissue provides crucial information about the functionality of the compounds and the tissue. This information is provided by hyperspectral imaging (HI) techniques, which associate spectral information with the pixel position in the sample. These techniques, however, generate a large amount of complex data that must be processed with chemometric tools, such as the Multivariate Curve Resolution-Alternating Least Squares (MCR-ALS) method, to enable a proper interpretation of the measurement. In addition, the complementary information provided by different spectroscopic hyperspectral imaging platforms can be combined using suitable image fusion strategies to obtain a more complete description of the biological tissue. This paper seeks to show the potential of the combination of hyperspectral imaging and chemometric tools to gather information associated with biological tissues. For this purpose, a study performed on several rice leaf and root microsections using Raman and fluorescence image fusion is described.

Keywords: Hyperspectral imaging, image fusion, chemometrics, plant tissue.

Introducció

Quan s'estudia un teixit biològic, els esforços s'orienten principalment a resoldre dues qüestions. La primera és: quins compostos químics hi ha al teixit? El fet de trobar determinades molècules en un teixit ens dona informació química valuosa sobre la funcionalitat tant de la molècula trobada com de la zona del teixit estudiada. A una fulla, per exemple, hi trobem proteïnes amb diferents funcions, com poden ser dur a terme reaccions enzimàtiques im-

plicades en la fotosíntesi o exercir un paper més estructural, entre d'altres. Només amb la informació química, però, és difícil esbrinar de manera clara el paper de la molècula al teixit. Quin paper hi tenen les proteïnes, en una fulla? Les proteïnes són presents només a les zones verdes de la fulla? N'hi ha, a l'esquelet de la fulla? Com les podem diferenciar? Si agafem una fulla i l'analitzem amb tècniques analítiques que demanen fer un extracte en solució de la mostra, com ara la cromatografia de líquids amb detecció d'espectrometria de masses (LC-MS), aconseguirem un molt bon perfil proteic, però haurèrem esborrat completament la connexió entre la molècula i la zona del teixit en què es localitza. Saber on es troba concretament el compost químic al teixit ens dona indicacions sobre la seva funció. Podem dir, llavors, que la informació química és necessària, però no sempre suficient, per estudiar el rol dels compostos químics en un teixit biològic.

Correspondència: Adrián Gómez-Sánchez
Universitat de Barcelona. Facultat de Química. Departament d'Enginyeria Química i Química Analítica. Grup de Quimiometria
C. de Martí i Franquès, 1-11. 08028 Barcelona
A/e: adrian.gomez.sanchez@ub.edu

Així doncs, la segona pregunta que cal respondre és: on es localitzen els compostos químics al teixit? Saber on és un compost químic al teixit vegetal ens ajuda a formular hipòtesis (o confirmar-les) sobre la funció tant del compost químic com del teixit. Per exemple, si trobem unes proteïnes a l'esquelet de la fulla de la planta, podem relacionar ràpidament aquestes molècules amb funcions estructurals. Si les trobem a les zones verdes (on es desenvolupa la fotosíntesi), podem connectar-les amb la producció d'energia (vegeu la figura 1). La informació espacial és aleshores tan valuosa com la informació química. La relació que neix d'aquí, entre la molècula i la localització, és una peça clau per comprendre l'estructura i la funció dels teixits biològics.

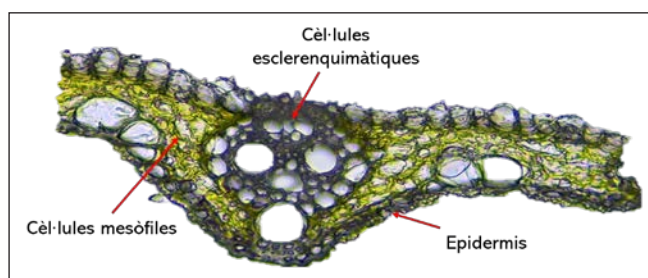


FIGURA 1. Imatge de microscopi òptic de la secció transversal d'una fulla d'arròs. S'hi poden veure les diferents zones de la fulla, com el sistema vascular, on es troben les cèl·lules esclerenquimàtiques, fortament lignificades, que donen robustesa a la fulla; les zones verdes, on es duu a terme la fotosíntesi a les cèl·lules mesòfiles, i, finalment, l'epidermis. Elaboració pròpia.

Per poder obtenir simultàniament la informació espacial i química d'un sistema, han sorgit tècniques analítiques com la imatgeria hiperespectral [1]. És una tècnica espectroscòpica que consisteix a fotografiar la mostra en diferents longituds d'ona. Com a resultat d'aquest procés, obtenim una imatge hiperespectral, que està formada per la superfície escanejada de la mostra dividida en píxels, de cadascun dels quals es mesura un espectre relacionat amb el material que conté. Així doncs, una imatge hiperespectral es visualitza molt clarament com un cub de dades de tres dimensions (3D), on dues dimensions són espacials (coordenades x i y dels píxels de l'àrea escanejada de la mostra) i una dimensió conté informació química (les diferents longituds d'ona) (figura 2).

Hi ha moltes imatges que responen a l'estructura en 3D que acabem d'esmentar, com ara les imatges Raman, d'infraroig (IR) o MS. També les imatges hiperespectrals de fluorescència més comunes s'adquireixen recollint l'espectre d'emissió associat a l'excitació dels píxels de la mostra mitjançant un làser discret, és a dir, d'una longitud d'ona concreta. No obstant

això, els avenços en microscòpia i espectrometria han permès el desenvolupament de nous microscopis de fluorescència, equipats amb làsers d'excitació d'espectre continu, que possibiliten treballar en un interval de longituds d'ona d'excitació. Això permet adquirir un nou tipus d'imatge hiperespectral, on a cada píxel li correspon una mesura en 2D d'excitació-emissió associada al material que conté. Així doncs, la imatge total obtinguda té quatre dimensions (4D), dues de les quals, espacials (les coordenades x i y dels píxels a l'àrea escanejada de la mostra), i dues d'informació química (associades a les longituds d'ona d'excitació i d'emissió) [2]. Aquesta dimensió extra és de gran ajuda quan els fluoròfors naturals dels teixits tenen un espectre d'emissió molt semblant entre ells i es fa difícil diferenciar-los. Si, a més de l'espectre d'emissió, tenim l'espectre d'excitació, la distinció de fluoròfors resulta més senzilla.

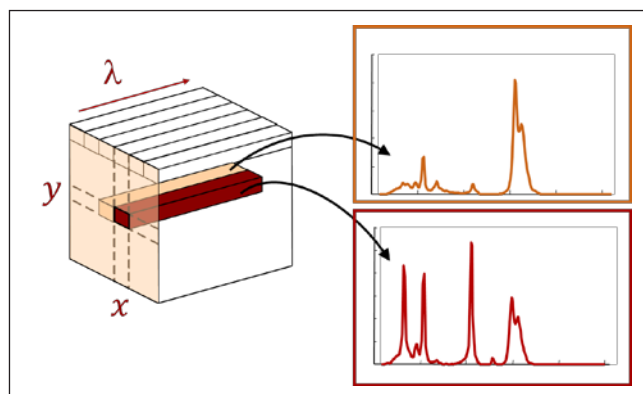


FIGURA 2. Esquema d'una imatge hiperespectral, formada per espectres associats als píxels de coordenades x i y . Elaboració pròpia.

Un altre aspecte rellevant és la possibilitat de combinar les dades procedents de diferents tècniques d'imatgeria hiperespectral i aprofitar la informació espectroscòpica complementària que presenten per caracteritzar millor morfològicament i químicament els components de la mostra. Aquesta aproximació es fa mitjançant estratègies de fusió d'imatges i s'aconsegueix combinar i tractar adequadament imatges en 3D que tenen diferents resolucions espacials [3] o imatges en 3D i 4D [2].

En els propers apartats, es mostrarà la utilitat de la imatgeria hiperespectral, en les modalitats d'imatges Raman, imatges en 3D i 4D de fluorescència, i es farà ús de la seva fusió per estudiar i descobrir els compostos químics que hi ha en seccions transversals de fulles de plantes d'arròs. La interpretació de la mesura es farà amb eines quimiomètriques, necessàries per gestionar la informació que contenen les dades de les imatges adquirides.

Les imatgeries hiperespectrals utilitzades

La imatgeria Raman és una de les tècniques usades en aquest treball. L'espectroscòpia Raman és una tècnica que consisteix a il·luminar la mostra amb un làser i recollir la dispersió de Raman. Els espectres Raman tenen una estructura molt fina i són molt útils per identificar molècules a partir de la seva empremta espectral. Les imatges Raman adquirides tenen una mida de píxel de 0,5 micres. En la figura 3c podem veure un exemple d'imatge Raman de la zona vascular d'una fulla, d'una llargària de 70 micres i una amplària de 20 micres, corresponent al senyal recollit en un desplaçament de Raman de $1\ 600\text{ cm}^{-1}$, propi d'una de les bandes de la lignina [4]. A la dreta es mostra l'espectre d'un píxel de la imatge, registrat en un interval de desplaçament de Raman de 270 fins a $2\ 015\text{ cm}^{-1}$.

Per complementar la mesura Raman, també s'han fet servir imatges hiperespectrals de fluorescència, que aprofiten la presència de fluoròfors naturals a les fulles. Una imatge hiperespectral de fluorescència s'obté excitant la mostra amb un làser

(discret en imatges en 3D i continu en imatges en 4D) i recollint l'emissió de fluorescència als diferents píxels. Les tècniques d'imatgeria de fluorescència permeten treballar amb una mida de píxel molt petita (fins i tot de desenes de nanòmetres). Això fa que sigui una tècnica molt atractiva, amb la qual és possible veure fins i tot cèl·lules individuals amb molt bona resolució espacial. La figura 3a mostra una imatge en 3D de fluorescència obtinguda utilitzant un làser d'excitació de 405 nm. La mida del píxel és de 200 nm i la imatge té unes dimensions de 190 micres de llargària i 190 micres d'amplària. L'interval espectral d'emissió recollit va de 420 fins a 750 nm. En la imatge mostrada, que correspon a la longitud d'ona d'emissió de 520 nm, hi podem veure clarament l'estructura de la fulla d'arròs: el sistema vascular central, format per cèl·lules estructurals (esclerenquimàtiques i col·lenquimàtiques), i les zones toves de la fulla, on se situen les cèl·lules encarregades de la fotosíntesi (mesòfiles) [5].

A diferència de l'espectroscòpia Raman, la informació química de les imatges de fluorescència no és tan específica, ja que els espectres tenen bandes amples i molt arrodonides i menys es-

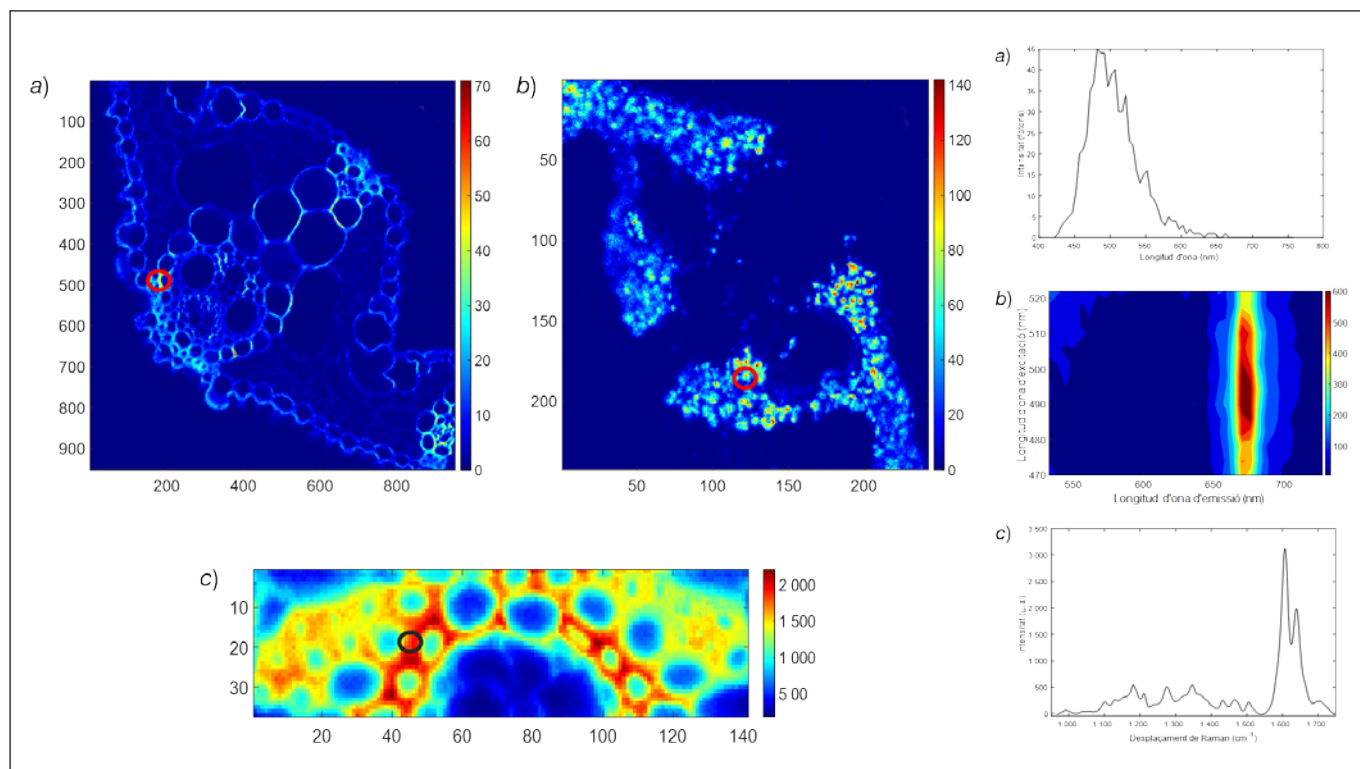


FIGURA 3. A l'esquerra: a) exemple d'una imatge de fluorescència en 3D d'una fulla de planta d'arròs, a una longitud d'ona d'excitació de 405 nm i d'emissió de 520 nm; b) exemple d'una imatge de fluorescència en 4D d'una fulla de planta d'arròs, a una longitud d'ona d'excitació de 490 nm i d'emissió de 650 nm; c) exemple d'una imatge Raman en 3D de la zona vascular d'una fulla d'arròs, a un desplaçament de Raman de $1\ 600\text{ cm}^{-1}$. A la dreta: espectres associats als píxels marcats a les imatges (a) espectre d'emissió de fluorescència; b) mapa en 2D d'excitació-emissió; c) espectre Raman). Elaboració pròpia.

tractura, com es pot observar a l'espectre d'emissió del píxel de la figura 3a. Una manera de compensar aquesta menor informació espectroscòpica és treballar amb la modalitat d'imatges de fluorescència en 4D, que, en afegir l'interval de longituds d'ona d'excitació, proporcionen una informació estructural més completa. En la figura 3b podem observar un exemple d'imatge de fluorescència en 4D, corresponent a una longitud d'ona d'excitació de 490 nm i d'emissió de 650 nm. Aquesta imatge té una mida de píxel de 700 nm, unes dimensions de 180 micres de llargària i amplària, i les mesures en 2D d'excitació-emissió de cada píxel cobreixen un interval d'excitació de 470 a 526 nm, i un interval d'emissió de 532,5 a 727,5 nm. A la dreta es representa el mapa en 2D d'excitació-emissió d'un píxel determinat.

La interpretació de les imatges hiperespectrals

Analitzar una imatge hiperespectral no és una qüestió trivial. Les imatges de fluorescència contenen milions d'espectres i , a més, amb informació poc selectiva. La situació no és gaire més simple en el cas de les imatges Raman. En ambdós casos, cal gestionar cubs o hipercubs de dades de grandària considerable i amb informació complexa.

La llei de Beer-Lambert. El model natural de la mesura de la imatge hiperespectral

Per triar eines útils de tractament de dades, cal conèixer el comportament natural de la mesura que es vol analitzar. Les imatges hiperespectrals estan formades per un gran conjunt d'espectres i , per tant, la llei de Beer-Lambert ens pot ser de gran ajuda com a model natural d'aquesta mesura.

Així, les imatges en 3D es poden entendre com un cub de dades i , alhora, aquest cub es pot desplegar com una taula o matriu de dades (D), on cada fila és l'espectre d'un píxel. El senyal dels espectres d'aquesta taula de dades es pot descriure com la suma de les contribucions al senyal (D) de cada compost present als píxels analitzats (equació 1). Cada contribució individual al senyal total es pot expressar com el producte de la concentració d'un determinat compost en cada píxel (c) per la seva signatura espectral pura (s_i^T) (equació 2). De manera més compacta, és possible expressar el senyal de tots els espectres de la taula D com a producte de la matriu C , que conté els perfils de concentració de tots els compostos presents en la imatge, per la matriu S^T , que conté les signatures espectrals pures que s'hi relacionen. L'equació 3 és el model bilineal que reflecteix la llei de Beer-Lambert i que podeu veure representat gràficament en la figura 4.

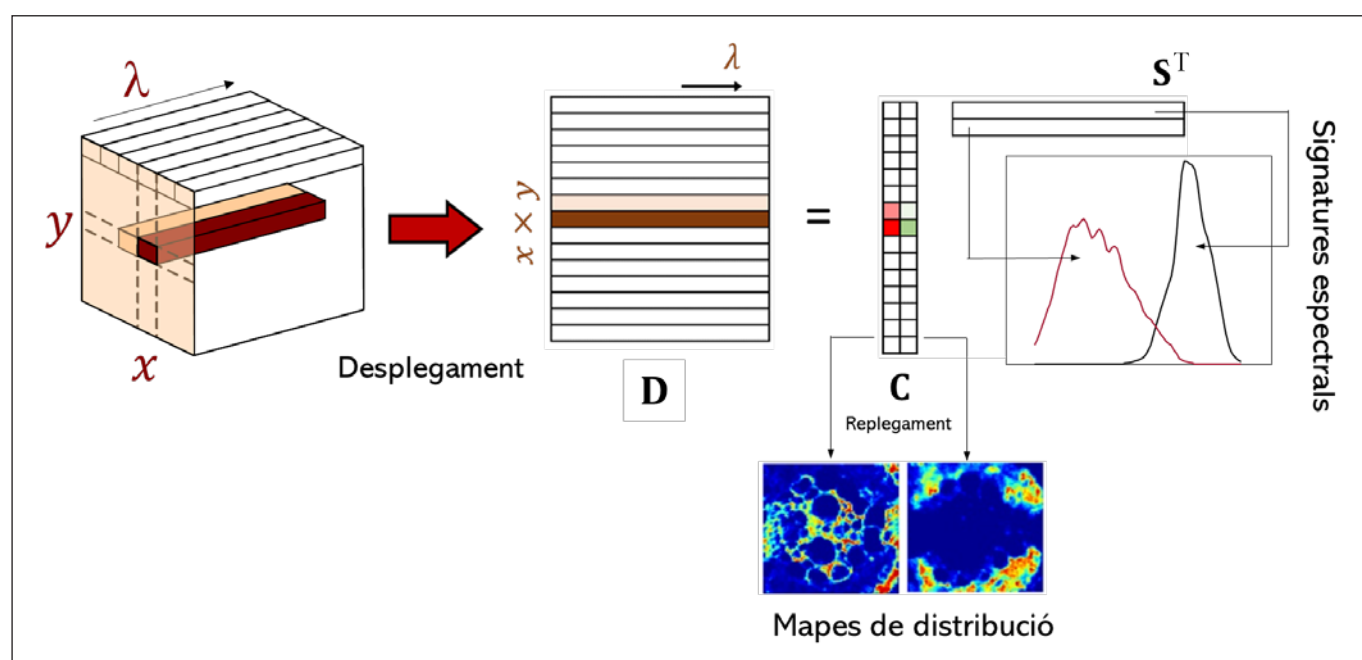


FIGURA 4. Esquema general del model bilineal propi d'una imatge hiperespectral en 3D. Elaboració pròpia.

$$\mathbf{D} = \sum_i \mathbf{D}_i + \quad (\text{eq. 1})$$

$$\mathbf{D} = \sum_i \mathbf{c}_i \mathbf{s}_i^T + \mathbf{E} \quad (\text{eq. 2})$$

$$\mathbf{D} = \mathbf{C}\mathbf{S}^T + \mathbf{E} \quad (\text{eq. 3})$$

Les signatures espectrals de la matriu \mathbf{S}^T descriuen la informació química de cada component de la imatge. Si volem recuperar la informació espacial de la imatge, cada perfil de concentració relacionat es pot replegar segons l'estructura en 2D original de la imatge i s'obtidran els mapes de distribució. És a dir, podrem saber on està situat el component a la mostra i en quina proporció.

Les imatges en 4D de fluorescència d'excitació-emissió no segueixen un model bilineal, sinó un model trilineal. Aquest model trilineal es pot entendre com una extensió de la llei de Beer-Lambert i es representa gràficament en la figura 5. En aquest cas, cada contribució d'un compost químic i ve representada per l'espectre d'emissió (\mathbf{s}_i^T), l'espectre d'excitació (\mathbf{k}_i) i el perfil de concentració (\mathbf{c}_i) [2].

La llei de Beer-Lambert proporciona un model matemàtic senzill, útil per interpretar les imatges hiperespectrals. Ara bé, experimentalment només disposem de la matriu \mathbf{D} dels espectres de la imatge. Com podem obtenir la informació de les

matrius \mathbf{C} i \mathbf{S} , en el cas de les imatges en 3D, i \mathbf{C} , \mathbf{K} i \mathbf{S} , en el cas d'imatges en 4D, només a partir de la imatge inicial?

Mètode de resolució multivariant de corbes per mínims quadrats alternats (MCR-ALS).

La recuperació del model de Beer-Lambert

Els mètodes de resolució multivariant són els encarregats d'obtenir les matrius \mathbf{C} i \mathbf{S} pròpies del model bilineal de Beer-Lambert a partir de la informació continguda en la matriu \mathbf{D} dels espectres de la imatge [6]. En certs camps d'aplicació de la imatge, com el dels sensors remots, aquests mateixos procediments reben el nom de *mètodes de desmesclament (unmixing methods)*, perquè el seu objectiu és, en efecte, recuperar els senyals i els mapes de distribució dels components purs a partir dels senyals mesclats propis dels espectres dels píxels de la imatge [7].

En aquest treball hem fet ús d'un algorisme de resolució multivariant de base matemàtica simple, molt potent i que s'adapta perfectament a la natura de la mesura espectroscòpica: el mètode de resolució multivariant de corbes per mínims quadrats alternats (MCR-ALS) [8, 9]. Aquest mètode recupera el model bilineal de Beer-Lambert associat als espectres de la imatge experimental \mathbf{D} seguint tres etapes:

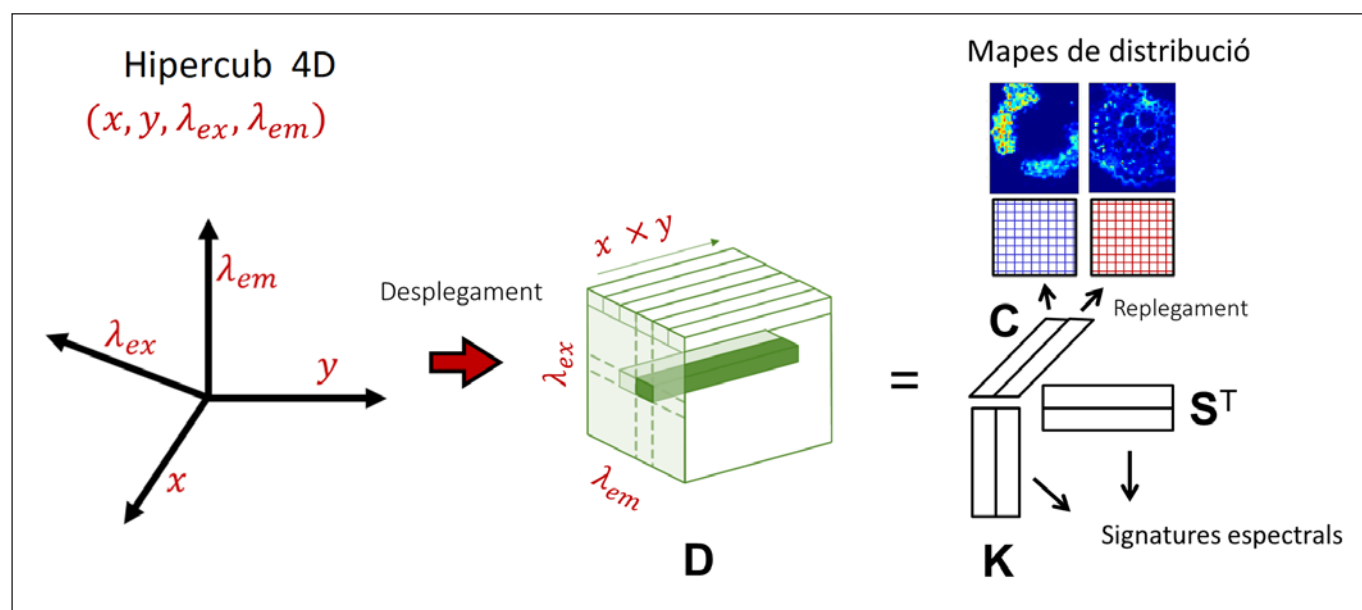


FIGURA 5. Esquema general del model trilineal propi d'una imatge hiperespectral en 4D de fluorescència d'excitació-emissió. Elaboració pròpia.

- 1) *Estimació del nombre de components químics presents a la mostra* i que es poden detectar amb la tècnica espectroscòpica utilitzada. Aquesta etapa es pot basar en el coneixement químic del teixit (o, en general, de la mostra estudiada). Des del punt de vista matemàtic, es poden usar mètodes quimiomètrics auxiliars, com l'anàlisi de components principals (PCA), que també descriu conjunts de dades mitjançant models bilineals de natura abstracta [10]. En general, el nombre de components principals necessari per descriure la matriu D en un model de PCA coincideix amb el nombre real de components de la imatge.
- 2) *Generació d'una estimació inicial de les matrius C o S^T* . Es poden construir a partir d'informació prèvia (per exemple, el coneixement d'algunes signatures espectrals pures, entre d'altres), però generalment s'utilitzen algorismes de selecció de variables pures, com el mètode *simple-to-use interactive self-modeling mixture analysis* (SIMPLISMA) [11], que permeten, a partir de les dades experimentals, seleccionar els espectres de la imatge que siguin més diferents

entre ells (S^T). Aquest tipus d'estimació inicial és un bon punt de partida per trobar les signatures espectrals pures dels components de la imatge.

- 3) *Optimització iterativa de les matrius C i S^T mitjançant un procediment de mínims quadrats alternats sota restriccions*. En cada cicle iteratiu, es calculen iterativament les matrius C i S fins que s'arriba a la convergència, segons el criteri que hàgim definit prèviament. En aquell punt, la solució correspon als perfils de concentració (C), que permetran recuperar els mapes de distribució dels components, i els espectres purs (S).

Durant el càlcul de l'etapa 3 es poden aplicar restriccions matemàtiques als perfils de les matrius C i S^T , que fan que les solucions obtingudes tinguin sentit químic [8, 9]. Per exemple, la restricció de no negativitat força els perfils de concentració i els espectres purs a no presentar valors negatius, per aconseguir una solució coherent amb la natura de la mesura, en què els valors negatius de concentracions, d'intensitats de

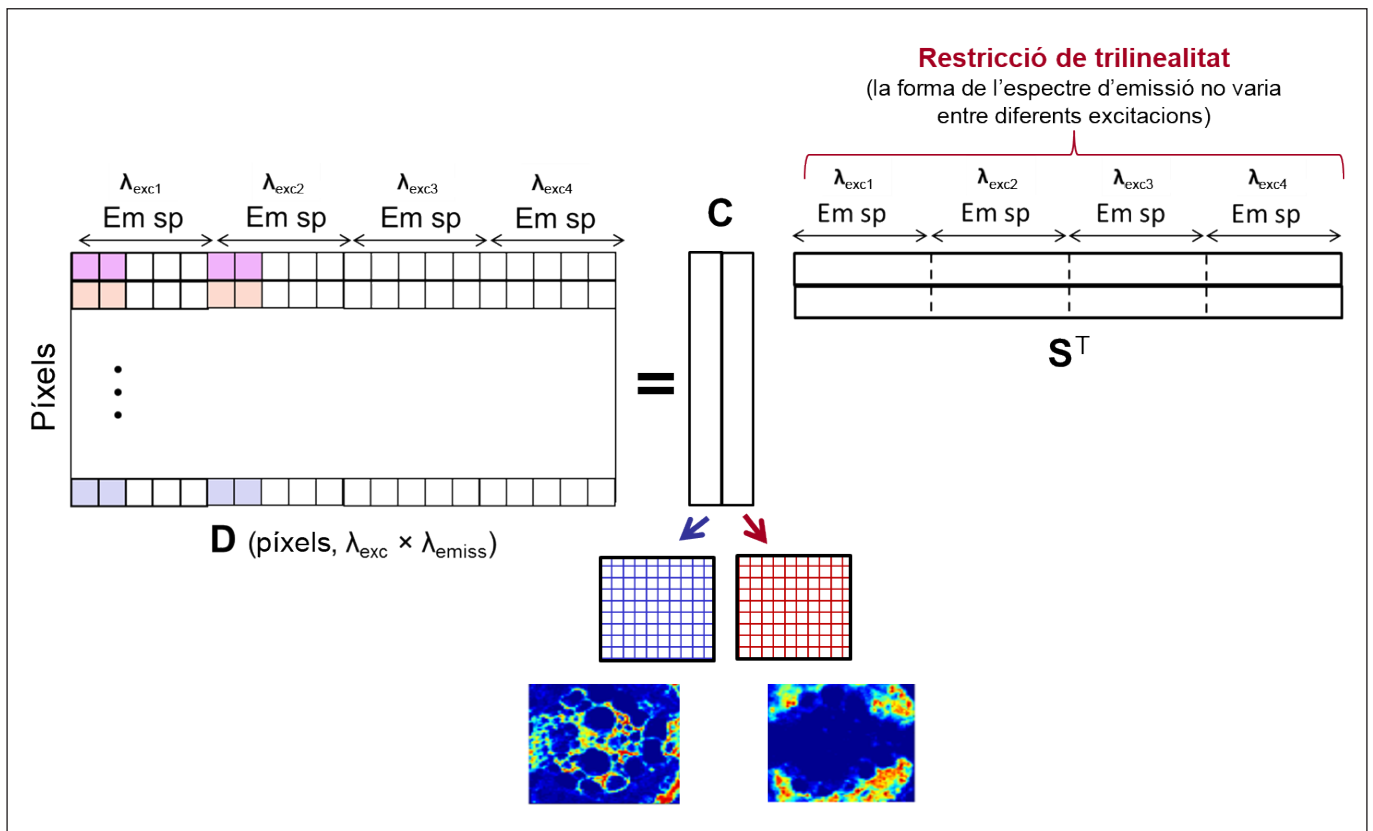


FIGURA 6. Transformació d'una imatge en 4D de fluorescència en una taula de dades. La imatge de quatre dimensions es desplega en forma de taula de dades concatenant en la mateixa fila tots els espectres d'emissió adquirits a diferents longituds d'ona d'excitació per a cada píxel. Elaboració pròpia.

fluorescència o d'intensitats Raman serien inacceptables. El mètode MCR-ALS permet implementar restriccions diferents a les matrius C i S i també als diferents components, si escau. Això converteix aquest algorisme en una eina molt potent i versàtil, capaç d'adaptar les restriccions a la natura de les tècniques emprades, al comportament dels components en les imatges analitzades i, en general, a multitud de sistemes químics d'altres tipologies [8, 9, 12, 13].

Per poder analitzar imatges en 4D, el mètode MCR-ALS permet, a través d'una restricció, aplicar un model trilineal en lloc d'un model bilineal i, així, respectar el model natural d'excitació-emissió de la mesura de les imatges hiperespectrals. Per fer-ho, la imatge en 4D s'ha de desplegar en forma de taula de dades concatenant a la mateixa fila tots els espectres d'emissió adquirits a diferents longituds d'ona d'excitació per a cada píxel, tal com es veu a la figura 6. La restricció de trilinealitat fa que els espectres d'emissió d'un component pur (és a dir, els blocs de cada fila de la matriu S^T de la figura 6) tinguin la mateixa forma i només variïn en intensitat [14]. En acabar l'anàlisi, es repleguen els perfils de concentració per obtenir els mapes de distribució. En el cas de la matriu S^T , es pot replegar cada perfil i mostrar la matriu en 2D d'excitació-emissió pura associada a cada component,

o simplement es poden representar les formes pures de l'espectre d'excitació i el d'emissió.

És important esmentar que el mètode MCR-ALS també pot treballar amb estructures multiconjunt (*multisets*), pròpies de la fusió d'imatges i que contenen blocs d'informació de diverses imatges relacionades [9, 14, 15]. En aquest cas, la matriu D es pot ampliar en la direcció de les columnes connectant espectres d'imatges hiperespectrals procedents de mostres mesurades mitjançant la mateixa tècnica espectroscòpica, sempre que l'interval espectral mesurat sigui comú per a totes elles. D'altra banda, les imatges adquirides sobre la mateixa mostra amb diferents plataformes espectroscòpiques també es poden fusionar ampliant la matriu D en la direcció de les files, sempre que els píxels tinguin la mateixa grandària i representin el mateix material en les imatges combinades. Respectant les condicions anteriors, es poden arribar a fusionar imatges captades sobre diferents mostres amb diferents plataformes espectroscòpiques. L'aplicació de restriccions en l'anàlisi d'estructures multiconjunt pren una importància especial perquè la tria de les restriccions pot canviar també segons els diferents blocs de les matrius C i S , lligats a diferents mostres o a diferents plataformes espectroscòpiques, respectivament [9, 14].

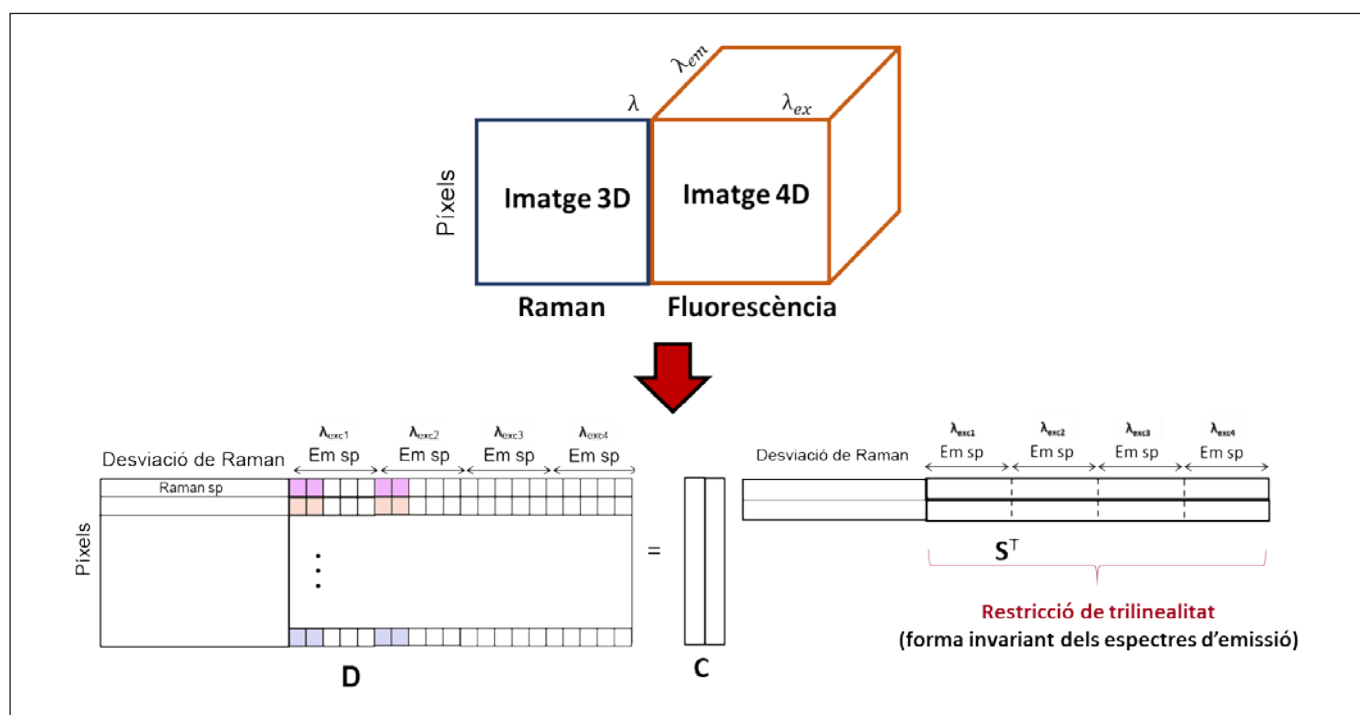


FIGURA 7. Esquema d'aplicació del mètode MCR-ALS per a la fusió d'imatges en 3D i 4D a partir de l'ús d'un model híbrid bilineal-trilineal. Elaboració pròpia.

Dins de la fusió d'imatges, s'ha aconseguit resoldre reptes importants utilitzant modificacions de l'algorisme MCR-ALS, que han permès treballar amb imatges de plataformes espectroscòpiques amb una resolució espacial diferent [3], o que han facilitat la gestió de la fusió d'imatges en 3D i 4D preservant els models bilineals i trilineals propis de cadascuna de les mesures [2]. La idea d'aquesta darrera variant està relacionada amb l'estudi de les fulles d'arròs mitjançant imatges en 3D Raman i en 4D de fluorescència; se'n mostra el funcionament d'una manera esquemàtica en la figura 7.

Com es pot veure a la figura 7, el primer pas consisteix a construir una estructura multiconjunt on es concatenen els espectres de la imatge en 3D Raman i en 4D de fluorescència (aquesta darrera, transformada en forma de taula, tal com ja s'ha explicat per a la figura 6). Per poder fer aquesta fusió, s'ha procedit a un preprocessament previ que ha igualat la mida de píxel de les dues tècniques i l'orientació espacial de les imatges adquirides [16]. Un cop es té la matriu D , s'aprofita el fet que les restriccions poden actuar en estructures multiconjunt per blocs i, així, només els blocs corresponents als espectres de fluorescència d'emissió de la matriu S^T són sotmesos a la restricció de trilinealitat, mentre que el bloc spectral de la imatge Raman queda exempt d'aquesta restricció [2]. Aquesta estratègia fa que, tot i ser a la mateixa estructura fusionada, cada imatge es descriu amb el model que és natural a la seva mesura (bilineal, en el cas de la imatge en 3D Raman, i trilineal, per a la imatge en 4D de fluorescència).

Els resultats de l'estudi de les fulles d'arròs. Imatges en 3D, 4D i fusió 3D-4D

La interpretació d'imatges hiperespectrals en 3D de fulles d'arròs

A continuació es mostraran dos exemples d'imatges hiperespectrals en 3D analitzades amb MCR-ALS: la imatge en 3D de fluorescència i la de Raman que s'han mostrat a les figures 3a i 3c, respectivament. Aquestes imatges tenen espectres de molta qualitat i han permès elucidar components amb diferències morfològiques i/o espectrals molt petites entre ells. Per a cada anàlisi, MCR-ALS dona com a resultat els mapes de distribució dels components amb els seus respectius espectres.

La figura 8 mostra els resultats de l'anàlisi de la imatge hiperespectral de fluorescència de la secció transversal d'una fulla d'arròs de la figura 3a. S'hi van poder detectar quatre components diferents i la variància explicada pel model bilineal va ser del 97,0%. Dos dels components, representats en vermell i groc, corresponen a diferents tipus de lignina [5]. La lignina és un component essencial a l'estructura de les plantes: els dona rigidesa i està situada principalment al sistema vascular. Tot i que les signatures espectrals d'aquests dos components són similars i només mostren un petit desplaçament entre elles, i que els seus mapes de distribució són mor-

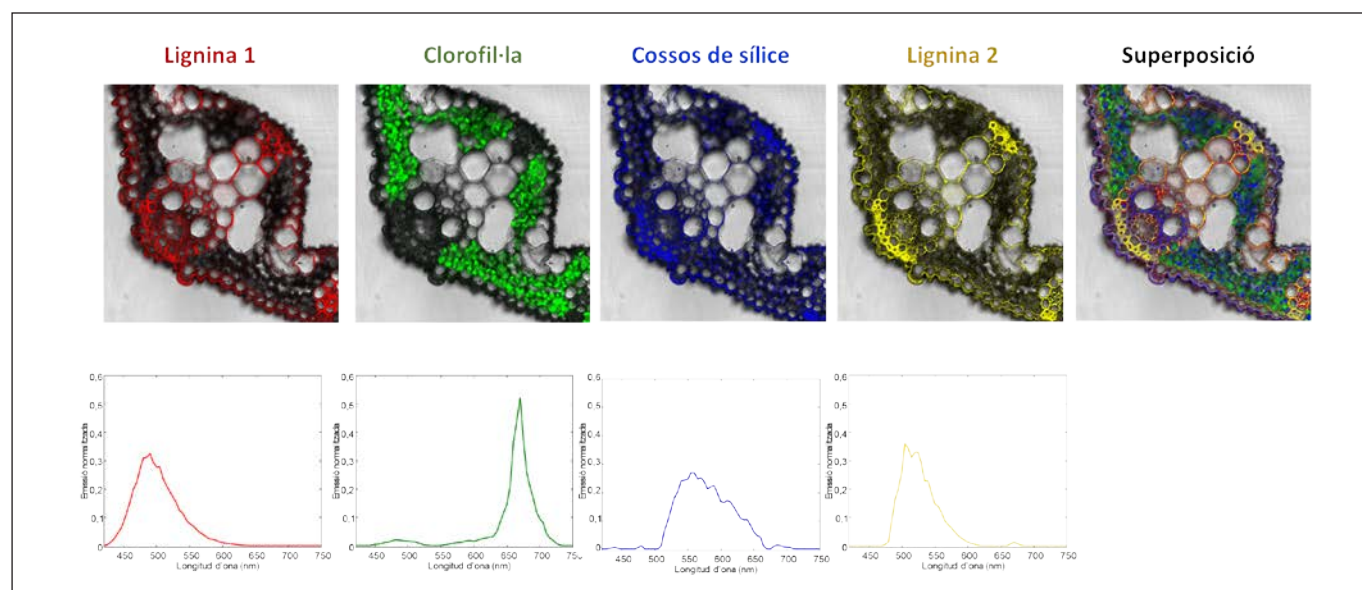


FIGURA 8. Mapes de distribució i signatures espectrals pures dels fluoròfors d'una imatge en 3D de fluorescència d'una fulla d'arròs. Elaboració pròpia.

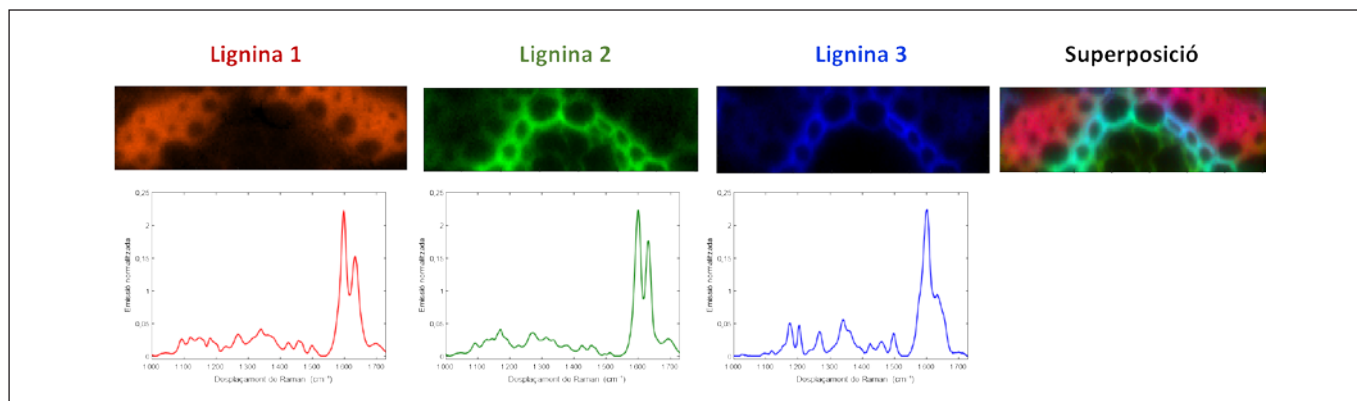


FIGURA 9. Mapes de distribució i signatures espectrals pures de compostos detectats amb la imatge en 3D Raman del sistema vascular de la fulla d'arròs. Elaboració pròpia.

foliògicament semblants, s'arriben a distingir, gràcies a la qualitat dels espectres inicials i a la gran resolució espacial de la tècnica. El component en blau apareix localitzat a la cutícula de la fulla i en petites acumulacions. Aquest component podria estar relacionat amb cossos de sílice, que ajuden a defensar la planta de possibles atacs [17]. Finalment, en verd, podem veure el component de la clorofil·la, amb un màxim d'emissió al voltant de 680 nm, definit en la bibliografia, i situat als cloroplasts [18]. El mapa en què estan representats tots els components dona idea de la potència de MCR-ALS, que aconsegueix destriar la contribució de cada component, tot i la seva superposició espacial i, en alguns casos, la similitud entre les signatures espectrals.

El segon exemple és de la imatge hiperespectral Raman de la zona vascular d'una secció transversal de fulla d'arròs, presentada en la figura 3c. Aquesta imatge és petita i molt localitzada a la zona vascular, i els espectres es van mesurar en condicions que van proporcionar una relació senyal/soroll molt elevada. En aquest cas, l'anàlisi per MCR-ALS hi va detectar tres components diferents, tal com es mostra en la figura 9, i el model bilineal va proporcionar una variància explicada del 99,98%. L'espectre Raman és molt selectiu i dona indicacions sobre els grups funcionals del compost. Com hem esmentat abans, la lignina és present al sistema vascular de les plantes. Podem veure que els tres components tenen un fort senyal de lignina (les dues bandes, prop de 1600 cm^{-1}) [4]. Hi ha, però, petites diferències en les signatures espectrals pures. La lignina està composta per tres monòmers diferents. A més, pot estar associada a diferents sucres, com la cel·lulosa. Aquestes diferències als espectres poden ser degudes a diferents proporcions dels monòmers constituents de la lignina o a la presència de diferents sucres associats. Es pot observar

com a les zones exteriors del sistema vascular hi predomina una lignina (en vermell), mentre que a l'interior hi predominen les formes representades en blau i verd.

La interpretació d'imatges hiperespectrals en 4D de fulles d'arròs

L'exemple següent és l'anàlisi de la imatge de fluorescència en 4D presentada en la figura 3b. Aquesta imatge està formada per intensitats de fluorescència mesurades en un interval de longituds d'ona d'excitació i d'emissió. En aquest cas, l'adquisició de la mesura és molt llarga i els espectres adquirits no tenen la mateixa qualitat que en el cas de la imatge en 3D de fluorescència mostrada anteriorment. A més, l'interval de longituds d'ona d'emissió també comença en longituds d'ona més elevades i la mostra és diferent. Per fer l'anàlisi MCR-ALS de la imatge, l'hipercub es va desplegar tal com s'ha esmentat anteriorment i es va aplicar la restricció de trilinealitat. En aquest cas, l'anàlisi per MCR-ALS hi va detectar dos components, com es veu en la figura 10, i el model va proporcionar una variància explicada del 99,34%. Els components identificats són la clorofil·la i la lignina, dels quals es presenten els espectres d'excitació i emissió. La posició dels dos components a la fulla és coherent amb la descripció esmentada anteriorment.

La fusió d'imatges en 3D Raman i en 4D de fluorescència en fulles d'arròs

La fusió d'imatges de diferents plataformes permet utilitzar la informació complementària de les tècniques per desmes-

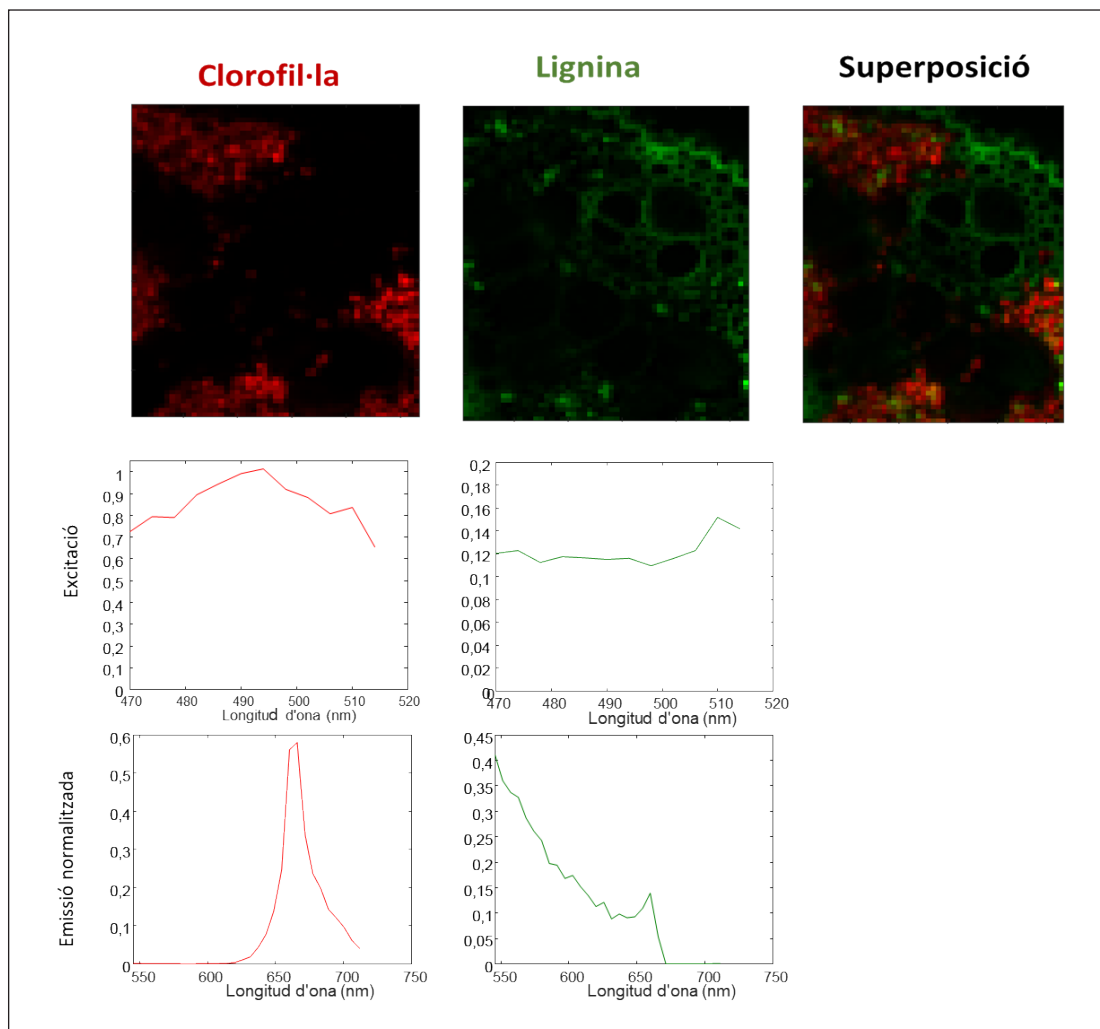


FIGURA 10. Mapes de distribució i signatures espectrals pures d'excitació i emissió dels fluoròfors de la fulla d'arròs. Elaboració pròpia.

clar millor els components de la mostra. Com hem descrit abans, la imatge en 3D Raman i la imatge en 4D de fluorescència responen a un model bilineal i a un model trilineal, respectivament. Per solucionar aquest problema, s'ha usat MCR-ALS amb un model híbrid que conserva la natura bilineal i trilineal de les dades que fusionem [2], tal com es mostrava en la figura 7.

En l'exemple que il·lustra la fusió, es van utilitzar imatges en 4D de fluorescència i imatges en 3D Raman d'una secció transversal que cobreix completament un vas secundari de la fulla d'arròs. L'anàlisi amb MCR-ALS distingeix la presència de tres components i el model té una variància explicada del 98,98%. La figura 11 mostra els mapes dels diferents components detectats i els seus respectius espectres purs. Com s'hi

observa, les dues tècniques utilitzades no detecten els mateixos components; per això, la fusió d'imatges és important per a una descripció més integral de la mostra. Així, es pot veure com la clorofil·la no té senyal Raman i el β -carotè no té fluorescència. Per aquest motiu, no se'n presenten espectres purs en les tècniques en què els components són indetectables. En canvi, la lignina mostra fluorescència i senyal Raman. La lignina i la clorofil·la presenten les característiques morfològiques i químiques esmentades anteriorment. El β -carotè s'identifica de manera inconfusible per les dues bandes Raman situades en desplaçaments de Raman d'entre 1 150 i 1 520 cm^{-1} , i perquè se situa a la zona pigmentada de la fulla, juntament amb la clorofil·la [19]. Com en altres casos, és la signatura espectral i la localització a la mostra el que permet identificar els components resultants.

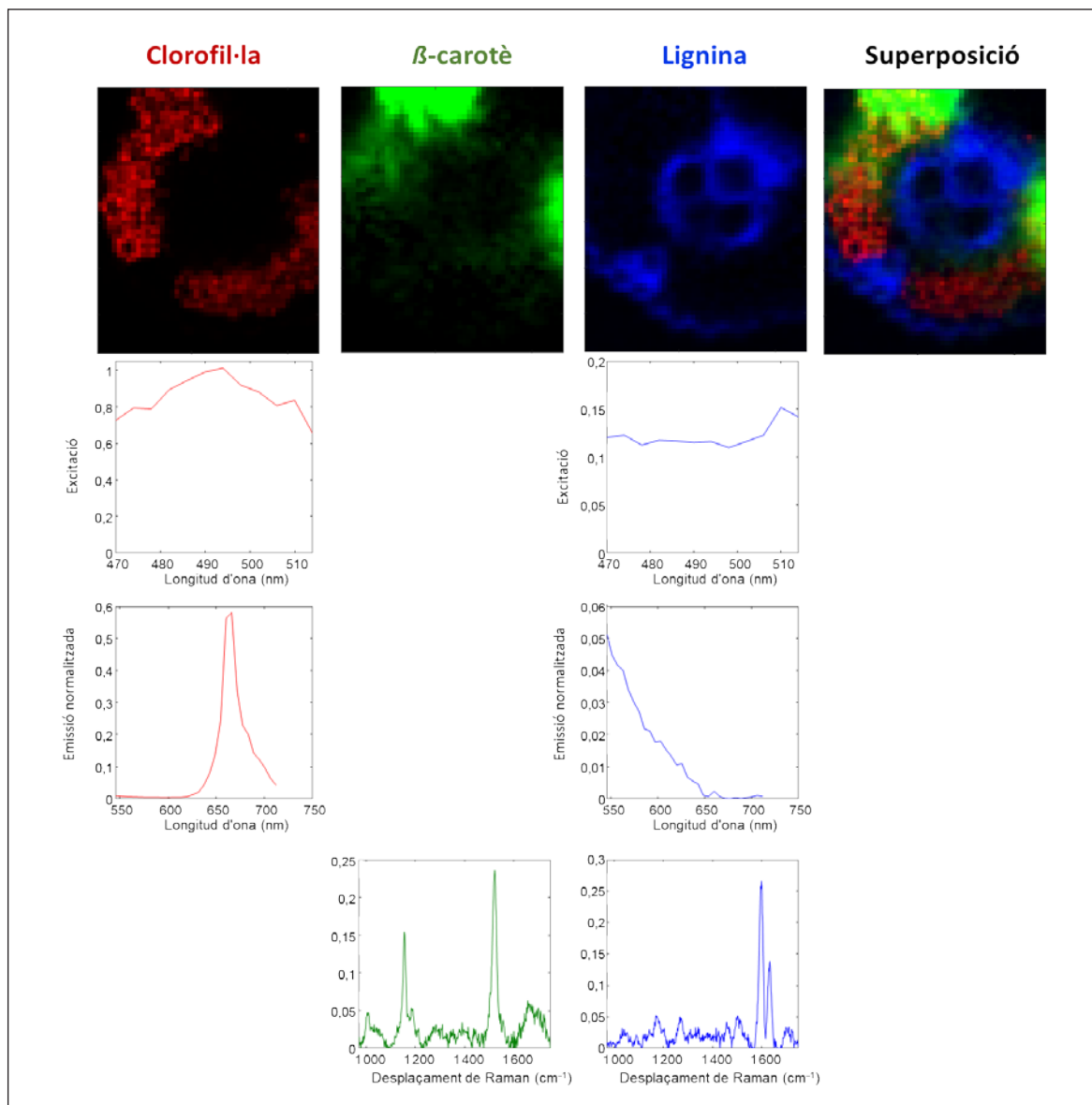


FIGURA 11. Mapes de distribució i signatures espectrals pures d'excitació, emissió i Raman de la secció d'un vas secundari d'una fulla d'arròs. Elaboració pròpia.

Conclusions

En resum, podem dir que les tècniques d'imatgeria hiperespectral combinades amb mètodes d'anàlisi de dades de resolució multivariant són eines excel·lents per caracteritzar morfològicament i químicament els compostos presents en mostres biològiques. En aquest sentit, quan diferents plataformes espectroscòpiques es combinen mitjançant estratègies adequades de fusió de dades, la informació complementària que proporcionen millora substancialment la capacitat de diferenciar components i la descripció integral de les mostres.

En l'ús de la imatgeria hiperespectral per estudiar teixits biològics s'han aconseguit fites importants, però resten fronts oberts connectats a la millora de les tècniques d'imatge i dels procediments de mesura, per fer-los menys fotolèsius i per alterar mínimament les mostres estudiades, i també al disseny de tractaments de dades que resolguin situacions en què la fusió d'imatges connecti plataformes amb diferents dimensions espacials (imatges en un pla i multicapa), o imatges estàtiques amb més resolució espectral i imatges dinàmiques (*light sheet*) amb poques bandes espectrals sobre mostres relacionades. Aquests i altres reptes fan que aquest camp d'investigació ofereixi un llarg i apassionant camí per recórrer.

Agraïments

Adrián Gómez-Sánchez i Anna de Juan agraeixen el suport econòmic del Ministeri de Ciència i Innovació, lligat al projecte PID 2019-1071586B-I00. Mónica Marro, María Marsal i Pablo Loza-Álvarez agraeixen el suport econòmic del Ministeri d'Economia i Competitivitat, mitjançant la beca AEI/FEDER; de la Fundació Privada Cellex; de la Fundació Privada Mir-Puig, i de la Generalitat de Catalunya, mitjançant el programa CERCA i Laserlab-Europe (acord núm. 871124).

Referències

- [1] AMIGO, J. M. (ed.). *Hyperspectral imaging*. 1a ed. Oxford: Elsevier, 2020. (Data Handling in Science and Technology; 32)
- [2] GÓMEZ-SÁNCHEZ, A.; MARRO, M.; MARSAL, M.; LOZA-ÁLVAREZ, P.; JUAN, A. de. «3D and 4D image fusion: Coping with differences in spectroscopic modes among hyperspectral images». *Anal. Chem.*, 92 (14) (2020), p. 9591-9602. DOI: 10.1021/acs.analchem.0c00780.
- [3] PIQUERAS, S.; BEDIA, C.; BELEITES, C.; KRAFFT, C.; POPP, J.; MAEDER, M.; TAULER, R.; JUAN, A. de. «Handling different spatial resolutions in image fusion by multivariate curve resolution: Alternating least squares for incomplete image multisets». *Anal. Chem.*, 90 (11) (2018), p. 6757-6765.
- [4] ZHANG, X.; CHEN, S.; XU, F. «Combining Raman imaging and multivariate analysis to visualize lignin, cellulose, and hemicellulose in the plant cell wall». *J. Vis. Exp.*, 124 (2017), e55910.
- [5] DONALDSON, L. «Softwood and hardwood lignin fluorescence spectra of wood cell walls in different mounting media». *IAWA J.*, 34 (1) (2013), p. 3-19.
- [6] RUTAN, S. C.; JUAN, A. de; TAULER, R. «Introduction to multivariate curve resolution: Iterative resolution methods». A: BROWN, S. D.; TAULER, R.; WALCZAK, B. (ed.). *Comprehensive chemometrics*. Vol. 2. 2a ed. Oxford: Elsevier, 2020, p. 85-94.
- [7] BIOCAS-DIAS, J. M.; PLAZA, A.; DOBIGEON, N.; PARENTE, M.; DU, Q.; GADER, P.; CHANUSSOT, J. «Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches». *IEEE JSTARS*, 5 (2) (2012), p. 354-379.
- [8] JUAN, A. de; RUTAN, S. C.; TAULER, R. «Two-way data analysis: Multivariate curve resolution: Iterative resolution methods». A: BROWN, S. D.; TAULER, R.; WALCZAK, B. (ed.). *Comprehensive chemometrics*. Vol. 2. 2a ed. Oxford: Elsevier, 2020, p. 153-171.
- [9] JUAN, A. de. «Multivariate curve resolution for hyperspectral image analysis». A: AMIGO, J. M. (ed.). *Hyperspectral imaging*. 1a ed. Oxford: Elsevier, 2020. (Data Handling in Science and Technology; 32), p. 115-150.
- [10] JOLIFFE, I. T. *Principal component analysis*. Springer, 1986.
- [11] WINDIG, W.; GUILMENT, J. «Interactive self-modeling mixture analysis». *Anal. Chem.*, 63 (1991), p. 1425-1432.
- [12] HUGELIER, S.; DEVOS, O.; RUCKEBUSCH, C. «On the implementation of spatial constraints in multivariate curve resolution alternating least squares for hyperspectral image analysis». *J. Chemom.*, 29 (2015), p. 557-561.
- [13] HUGELIER, S.; PIQUERAS, S.; BEDIA, C.; JUAN, A. de; RUCKEBUSCH, C. «Application of a sparseness constraint in multivariate curve resolution-alternating least squares». *Anal. Chim. Acta*, 1000 (2018), p. 100-108.
- [14] TAULER, R.; MAEDER, M.; JUAN, A. de. «Multiset data analysis: Extended multivariate curve resolution». A: BROWN, S. D.; TAULER, R.; WALCZAK, B. (ed.). *Comprehensive chemometrics*. Vol. 2. 2a ed. Oxford: Elsevier, 2020, p. 305-336.
- [15] JUAN, A. de; GOWEN, A.; DUPONCHEL, L.; RUCKEBUSCH, C. «Image fusion». A: COCCHI, M. *Data fusion methodology and applications*. Amsterdam: Elsevier, 2019, p. 311-344.
- [16] PIQUERAS, S.; MAEDER, M.; TAULER, R.; JUAN, A. de. «A new matching image preprocessing for image data fusion». *Chemom. Intell. Lab. Sys.*, 164 (2017), p. 32-42.
- [17] AGARIE, S.; AGATA, W.; UCHIDA, H.; KUBOTA, F.; KAUFMAN, P. B. «Function of silica bodies in the epidermal system of rice (*Oryza sativa* L.): Testing the window hypothesis». *J. Exp.*, 47 (1996), p. 655-660.
- [18] NDAO, A. S.; KONTÉ, A.; BIAÏE, M.; FAYE, M. E.; FAYE, N. A. B.; WAGUÉ, A. «Analysis of chlorophyll fluorescence spectra in some tropical plants». *J. Fluoresc.*, 15 (2) (2005), p. 123-129.
- [19] TSCHIRNER, N.; SCHENDERLEIN, M.; BROSE, K.; SCHLODDER, E.; MROGINSKI, M. A.; THOMSEN, C.; HILDEBRANDT, P. «Resonance Raman spectra of β -carotene in solution and in photosystems revisited: An experimental and theoretical study». *Phys. Chem. Chem. Phys.*, 11 (2009), p. 11471-11478.



A. Gómez-Sánchez



M. Marro



M. Marsal



P. Loza-Álvarez



A. de Juan

Adrián Gómez-Sánchez és graduat en química per la Universitat de Barcelona (UB). Actualment és estudiant de doctorat al Grup de Quimiometria de la UB. La seva tesi doctoral té com a objectiu el desenvolupament i l'aplicació de noves tècniques de fusió de dades en el camp de la imatgeria hiperespectral.

Mónica Marro és doctora en fotònica i biofotònica per la Universitat Politècnica de Catalunya (UPC, 2013) i actualment forma part de la Unitat de Microscòpia de Superresolució i Nanoscòpia de l'Institut de Ciències Fotòniques (ICFO). La seva recerca se centra a explotar l'espectroscòpia Raman en un ampli rang d'aplicacions biomèdiques: des de les molècules individuals fins a les cèl·lules i els teixits.

María Marsal és doctora en biologia per la Universitat de Barcelona (UB, 2005) i actualment és investigadora postdoctoral a l'equip de Pablo Loza-Álvarez, a la Unitat de Microscòpia de Superresolució i Nanoscòpia de l'Institut de Ciències Fotòniques (ICFO). La seva recerca se centra en els avenços tècnics de la microscòpia de superresolució i confocal, en estudis d'imatge de calci i en estudis del desenvolupament del peix zebra a partir de tècniques d'imatge.

Pablo Loza-Álvarez es va doctorar en física de làsers per la Universitat de Saint Andrews (Escòcia, Regne Unit) l'any 2000 i la seva investigació se centra en el desenvolupament de noves tècniques de microscòpia que van des de la superresolució fins al nivell macroscòpic, i per a aplicacions en biomedicina. Actualment és el cap de la Unitat de Microscòpia de Superresolució i Nanoscòpia de l'Institut de Ciències Fotòniques (ICFO).

Anna de Juan és doctora en química per la Universitat de Barcelona (UB, 1997) i la seva recerca se centra en el desenvolupament i l'aplicació de mètodes quimiomètrics a imatges hiperespectrals i processos. Ha estat professora convidada a la Universitat de Mòdena i Reggio de l'Emília (Itàlia), a la Universitat Dalhousie (Canadà) i a la Universitat de Lilla I (França). Rebé el Chemometrics Elsevier Award el 2004. Pertany a l'equip editorial de les revistes *Analytica Chimica Acta* i *Chemometrics and Intelligent Laboratory Systems*. Actualment és professora titular de la Facultat de Química de la UB.