

УДК 535.338

КОРРЕКЦИЯ ОШИБОК В МАТРИЦАХ ВОЗБУЖДЕНИЯ-ЭМИССИИ С ПОМОЩЬЮ МЕТОДА ГЛАВНЫХ КОМПОНЕНТ

© Н. А. Маслов

*Институт теоретической и прикладной механики им. С. А. Христиановича СО РАН,
630090, г. Новосибирск, ул. Институтская, 4/1
E-mail: nmaslov@itam.nsc.ru*

Флуоресцентная спектроскопия, использующая сканирование длин волн возбуждения, представляет собой мощный химический, биологический и медицинский диагностический инструмент, позволяющий охарактеризовать распределение флуорофоров в образцах различной природы. Однако, когда флуорофоры изначально неизвестны, интерпретация больших наборов спектров флуоресценции, зависящих от длины волны возбуждения, значительно затрудняется, особенно если в измеренных матрицах возбуждения-эмиссии присутствуют потери данных вследствие случайных причин или систематических факторов. Предлагается использование метода главных компонент для восстановления повреждённых участков спектра, который, благодаря независимости спектров флуоресценции отдельных флуорофоров от длины волны возбуждающего излучения, позволяет делать это без потери данных.

Ключевые слова: лазерно-индуцированная флуоресценция, матрица возбуждения-эмиссии, метод главных компонент.

DOI: 10.15372/AUT20250208

EDN: SHDMIY

Введение. Метод лазерно-индуцированной флуоресценции (ЛИФ) основан на испускании фотонов молекулами, возбуждаемыми лазерным излучением. В конденсированном веществе невозможно идентифицировать отдельные спектральные линии из-за перекрытия энергетических уровней, в результате чего спектры флуоресценции непрерывны. ЛИФ-спектроскопия применяется при исследовании газовых потоков и пламени [1], а также в различных химических [2] и медицинских приложениях [3, 4]. В большинстве случаев авторы публикаций ограничиваются наблюдением различий между спектрами отдельных образцов. Часто сравнение результатов работ разных авторов затруднено, поскольку в их экспериментах использовались различные длины волн возбуждения.

Зависимость спектра флуоресценции от длины волны возбуждения (матрица возбуждения-эмиссии) позволяет получить больше спектральной информации, которую можно использовать для последующего анализа состава образца, поскольку каждый флуорофор имеет свои собственные уникальные спектры возбуждения и флуоресценции. В большинстве работ авторы проводят только визуальный анализ наблюдаемых спектров для описания состава образца и динамики флуоресценции, который обычно основан на информации о максимумах флуоресценции хорошо известных флуорофоров. Спектр образца часто представляется как сумма спектров известных флуорофоров с использованием различных моделей распространения излучения [5]. Однако такой анализ не всегда возможен, так как некоторые флуорофоры могут изменять свой спектр из-за взаимодействия с окружающей средой, а в образце могут присутствовать изначально неизвестные флуорофоры. Существуют различные способы анализа, основанные на статистических различиях матриц, полученных от разных образцов: метод главных компонент (ГК) [6], анализ многомерных кривых [7], параллельный факторный анализ [8]. Все они не требуют наличия априорной

информации о спектрах чистых флуорофоров. Однако в матрицах возбуждения-эмиссии могут присутствовать артефакты, не имеющие отношения к флуоресценции, такие как сигналы от источника возбуждающего излучения, комбинационного рассеяния, случайные засветки. Без информации о присутствующих в образце флуорофорах отделить полезный сигнал от паразитного становится проблематично.

Одним из способов очистки данных для разных приложений является использование метода главных компонент в различных вариациях, но, как правило, в таких случаях речь идёт о восстановлении данных с потерями [9].

Цель работы — представить способ очистки матриц возбуждения-эмиссии флуоресценции, основанный на том, что они представляют собой набор коррелирующих кривых, поскольку в большинстве случаев формы спектров флуоресценции отдельных флуорофоров не зависят от длины волны возбуждающего излучения. Таким образом, можно заместить повреждённые участки спектров подходящей линейной комбинацией главных компонент без потери значимой информации.

Материалы и методы.

Метод главных компонент. В сложных многокомпонентных образцах матрицы возбуждения-эмиссии могут быть описаны следующим образом [5]. Пусть образец освещается лазерным излучением с длиной волны $\lambda^{(ex)}$, а на длине волны $\lambda^{(fl)}$ регистрируется интенсивность его флуоресценции $F(\lambda^{(ex)}, \lambda^{(fl)})$. Существует широкий круг задач в биологии, медицине, анализе материалов, когда флуорофоры в образце находятся в твёрдом или жидкостном окружении или в составе большой молекулы (полимера). В этом случае из-за взаимодействия с окружением и, как следствие, быстрой колебательной релаксации возбуждённого электронного состояния для отдельного флуорофора номер i его нормированный спектр флуоресценции $f_i(\lambda^{(fl)})$ не зависит от длины волны возбуждающего излучения и определяется только электронно-колебательной структурой уровней. Эффективность возбуждения флуорофора зависит от его структуры, а также особенностей распространения и поглощения лазерного излучения в веществе образца и в результате от длины волны возбуждения $e_i(\lambda^{(ex)})$. Соответственно, вклад флуорофора номер i в итоговый сигнал будет выражаться произведением $e_i(\lambda^{(ex)})f_i(\lambda^{(fl)})$. Без учёта нелинейных эффектов, таких как реабсорбция флуоресценции, полную интенсивность можно представить как сумму вкладов всех N флуорофоров. При этом следует учитывать, что в реальном эксперименте всегда присутствуют шумы, которые для каждого измерения со своими $\lambda^{(ex)}$ и $\lambda^{(fl)}$ внесут случайную добавку $X(\lambda^{(ex)}, \lambda^{(fl)})$. Таким образом, зарегистрированная интенсивность будет равна

$$F(\lambda^{(ex)}, \lambda^{(fl)}) = \sum_{i=1}^N e_i(\lambda^{(ex)})f_i(\lambda^{(fl)}) + X(\lambda^{(ex)}, \lambda^{(fl)}).$$

Если в процессе эксперимента длина волны возбуждения сканируется и принимает одно из M значений $\lambda_j^{(ex)}$, а спектрометр измеряет интенсивность флуоресценции на длинах волн $\lambda_k^{(fl)}$, полученные значения можно представить в виде матрицы данных $F_{jk} = F(\lambda_j^{(ex)}, \lambda_k^{(fl)})$, для которой действительно разложение

$$F_{jk} = \sum_{i=1}^N e_{ij}f_{ik} + X_{jk}, \quad (1)$$

где $e_{ij} = e_i(\lambda_j^{(ex)})$, $f_{ik} = f_i(\lambda_k^{(fl)})$ и $X_{jk} = X(\lambda_j^{(ex)}, \lambda_k^{(fl)})$. Как уже было отмечено выше, часто спектры отдельных флуорофоров нельзя считать заранее известными, при этом они

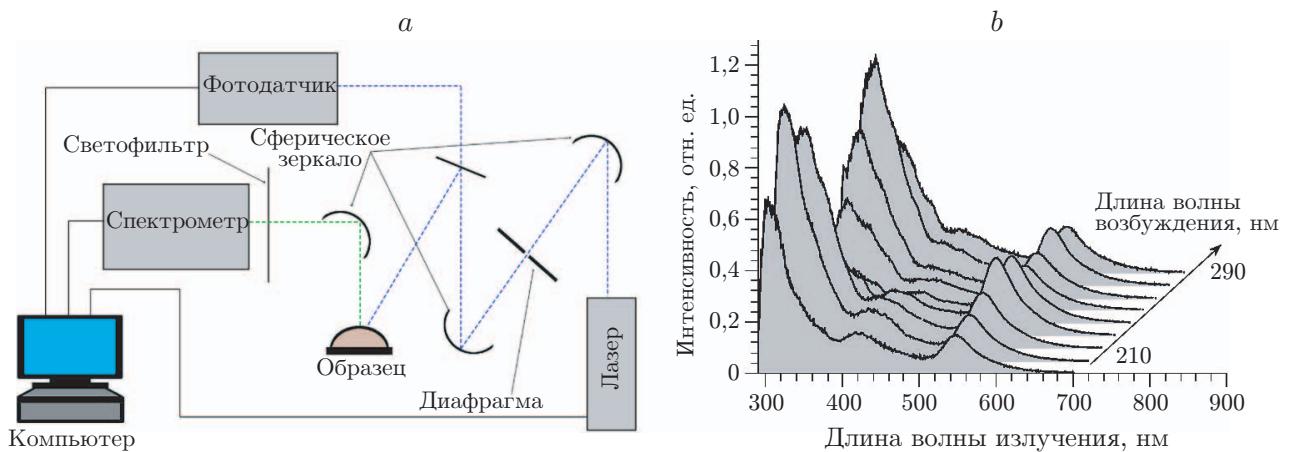


Рис. 1. Схема измерения спектров флуоресценции (а) и пример матрицы возбуждения-эмиссии (б)

могут существенно накладываться друг на друга. Поэтому приходится искать разложения вида (1) без априорной информации о флуорофорах. Хотя спектр каждого отдельного флуорофора не зависит от длины волны возбуждающего излучения, у всех флуорофоров разные зависимости эффективности возбуждения. Таким образом, суммарный спектр будет варьироваться при изменениях длины волны возбуждающего излучения, которые могут быть проанализированы статистическими методами. В отсутствие шумов задача сводится к поиску сингулярного разложения матрицы F_{jk} . При их наличии выполняется поиск наилучшего приближения, минимизирующего вклад шумов, — евклидовой нормы матрицы X_{jk} . К сожалению, однозначного разложения не существует, и оно определено с точностью до смены базиса векторов f_i . В зависимости от типа задачи применяются различные подходы. Самый простой из них — метод ГК (МГК), в котором последовательно осуществляется поиск ортогональных векторов f_i по мере уменьшения их вкладов в результирующую матрицу. Наиболее часто для этого используется нелинейный итеративный метод наименьших квадратов (NIPALS), последовательно находящий наилучшее приближение матрицы остатков [10]. Из-за ортогональности ГК их затруднительно связать со спектрами реальных флуорофоров, а ввиду неоднозначности разложения точное их определение представляет собой нетривиальную задачу [11]. Однако для нас важно, что с помощью МГК возможно с хорошей точностью найти разложение вида (1) и оценить количество ГК N , т. е. получить такое их число, чтобы матрица шумов была с достаточной для исследователя точностью мала. Если при этом окажется, что $M > N$, т. е. количество флуорофоров меньше количества измеренных спектров, это будет означать, что спектр, полученный при использовании любой длины волны возбуждения, можно представить в виде линейной комбинации главных компонент, вычисленных без его участия, чем мы воспользуемся для замещения повреждённых участков спектров.

Измерение спектров. Принципиальная схема экспериментальной установки представлена на рис. 1, а. В качестве источника возбуждения использовалась лазерная система VIBRANT (НЕ) 355 II + UV (OPOTEK, Inc., США) с оптическим параметрическим генератором (ОПО). Лазерный пучок оптически фильтровался с применением двух сферических зеркал и диафрагмы для отсечения рассеянного света от системы ОПО. Делитель оптического пучка направлял лазерное излучение на образец и фотодиод для измерения энергии импульса. Исследуемый образец помещали в плоскую кювету из нержавеющей стали. Размер лазерного пятна на поверхности образца составлял 5×10 мм. Сферическое зеркало направляло флуоресцентный сигнал на спектрометр Acton SpectraPro SP2300 с ПЗС-матрицей Pixis 256 (Princeton Instruments, США) для регистрации спектра образца. Фильтр высоких частот с пределом коротковолнового диапазона 300 нм расположен

перед входной щелью спектрометра для отсечения лазерного излучения с целью защиты спектрометра. Всё оборудование помещено в тёмный ящик и защищено от рассеянного света.

Длины волн возбуждения сканировали в диапазоне от 210 до 290 нм с шагом 10 нм. Спектры флуоресценции регистрировали с разрешением 0,5 нм для каждой длины волны возбуждения. В результате для всех исследованных образцов регистрировали спектральные данные в виде матриц возбуждения-эмиссии (рис. 1, *b*). Матрицы нормировали на максимум интенсивности и представляли в виде каскадной диаграммы, в которой длина волны флуоресценции находится на горизонтальной оси, интенсивность излучения — на вертикальной, а зависимость от длины волны возбуждения представлена диагональным сдвигом, где каждый спектр соответствует длине волны возбуждающего излучения от 210 до 290 нм.

Для демонстрации возможных артефактов измерений использовали наиболее удобные для этой цели образцы: пластик (неокрашенный полипропилен) и гомогенизированную ткань аорты коровы, разведённую физиологическим раствором в соотношении 1 : 1000. Как можно видеть (см. рис. 1, *b*), спектры действительно состоят из множества широких спектральных полос, которые накладываются друг на друга. При изменении длины волны возбуждающего излучения спектр ЛИФ меняется, что говорит о том, что в его формировании участвуют несколько флуорофоров, но из-за взаимного перекрытия их спектров визуально разделить их не представляется возможным.

Результаты.

Примеры артефактов измерений. К нарушению целостности данных матриц возбуждения-эмиссии флуоресценции могут приводить разнообразные причины. Это могут быть случайные, вовремя необнаруженные эффекты, такие как солнечная активность и воздействие космических частиц на матрицу фотодетектора, случайные засветки. В то же время возможны систематические ошибки, последствия которых нельзя устраниТЬ повторными измерениями. Авторам данной работы приходилось сталкиваться с описанными ниже эффектами.

Возможность перестраивать длину волны лазерного излучения в ультрафиолетовом диапазоне в описанной выше установке приводит к своим ограничениям. Поскольку ультрафиолетовое излучение генерируется в ней путём удвоения частоты излучения, полученного с помощью оптопараметрического преобразования, на выходе присутствуют два пучка: один с искомой длиной волны, другой — с излучением накачки с удвоенной длиной волны и на порядок большей интенсивностью. Так как длина волны обоих пучков меняется при перестройке, отделить один пучок от другого пассивным фильтром во всём диапазоне длин волн не представляется возможным. Для этой цели в конструкции лазера на выходе предусмотрен монохроматор, разделяющий пучки с помощью призмы. Однако высокая интенсивность исходного пучка приводит к тому, что рассеянное на элементах монохроматора излучение накачки распространяется в направлении рабочего пучка. Оптическая фильтрация (см. рис. 1) частично помогает решить эту проблему. Тем не менее, хотя интенсивность прошедшего излучения накачки в результате принятых мер существенно уменьшается, при регистрации слабых сигналов флуоресценции она может вносить существенный вклад в спектр (рис. 2). Здесь представлены примеры спектров полипропилена, полученные с использованием длин волн возбуждающего излучения 270 и 280 нм. Излучение накачки присутствует в виде пиков с максимумами на длинах волн 540 и 560 нм соответственно. Высокая интенсивность этих пиков приводит к тому, что значимый сигнал флуоресценции сравним по интенсивности с широким основанием пика, определяющимся аппаратной функцией дифракционной решётки.

Описанную выше проблему можно было бы решить подбором индивидуального фильтра для каждой длины волны возбуждающего излучения, однако для коротковолнового

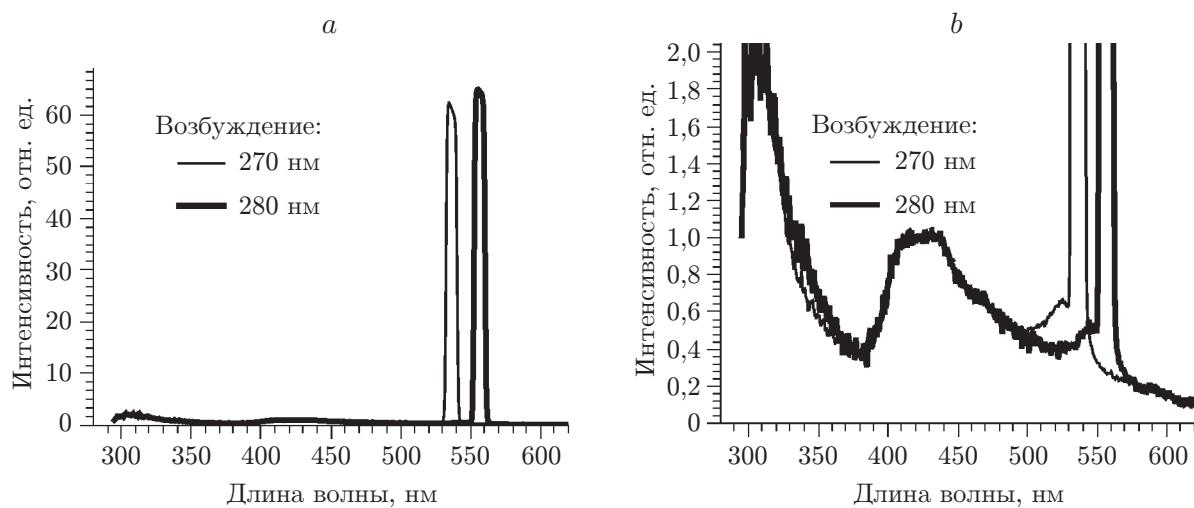


Рис. 2. Примеры влияния излучения накачки на регистрируемые спектры флуоресценции: *a* — исходный спектр полностью, *b* — увеличенный фрагмент

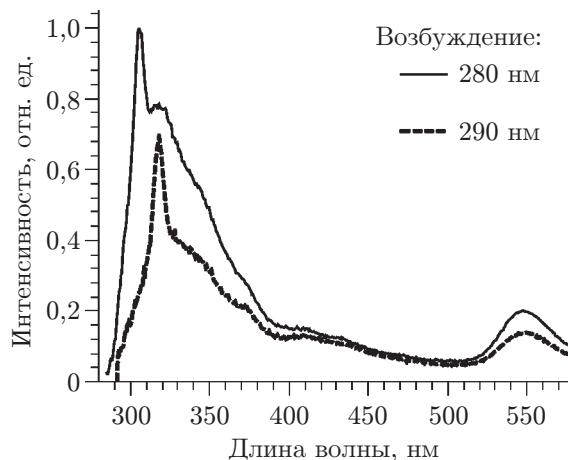


Рис. 3. Примеры влияния рассеяния на регистрируемые спектры

ультрафиолетового излучения это проблематично. Кроме того, возможны другие явления, от которых невозможно избавиться правильной подготовкой возбуждающего излучения.

При регистрации слабых сигналов на спектры флуоресценции могут также накладываться пики комбинационного рассеяния лазерного излучения. На рис. 3 представлены спектры супензии гомогенизированной ткани аорты коровы. При возбуждении флуоресценции излучением с длинами волн 280 и 290 нм в регистрируемых спектрах, помимо сигнала флуоресценции, присутствуют пики комбинационного рассеяния на длинах волн 306 и 320 нм соответственно. В этом случае отделить один сигнал от другого оптическими методами невозможно.

Коррекция ошибок. Как показывает опыт, в измеренных матрицах возбуждения-эмиссии могут присутствовать артефакты различной природы, однако общим для них является зависимость их спектра от длины волны возбуждающего излучения. Замену повреждённых участков спектра можно проводить, используя аппроксимацию удобной функцией, например полиномом подходящей степени. Недочётом таких методов является появление ошибок, связанных с недостаточно точным описанием реальных спектров выбранной функцией. Особенно такие ошибки возникают при наличии точек перегиба на повреждённом участке.

Однако, как было отмечено выше, с учётом природы флуоресценции возможно полное восстановление данных из повреждённых участков спектра, поскольку матрицы возбуждения-эмиссии являются набором коррелирующих данных. Каждый спектр для каждой длины волны возбуждающего излучения представляет собой линейную комбинацию спектров отдельных флуорофоров, присутствующих в образце. При этом, благодаря быстрой колебательной релаксации, спектр каждого из флуорофоров не зависит от длины волны возбуждающего излучения. Вместе с тем эффективность возбуждения каждого из них меняется с изменением длины волны возбуждающего излучения, чем и обусловлено изменение при этом общего суммарного спектра. Таким образом, независимо от длины волны возбуждения отдельного образца спектр будет представлять собой линейную комбинацию из одних и тех же компонент (1). Точное определение спектров чистых флуорофоров представляет собой нетривиальную задачу [11], однако определение описывающей все различия набора спектров может быть получено МГК. Оказалось, что общее количество компонент в исследованных матрицах, как правило, не превышает четырёх. При этом в диапазоне длин волн излучения, возбуждающего флуоресценцию этих компонент, возможно проведение измерений в большем количестве, чем количество компонент (в нашем случае — девяти, лазерное излучение изменяли в диапазоне 210–290 нм с шагом 10 нм). Таким образом, поведение спектра на повреждённом участке возможно восстановить, используя спектры, полученные с помощью других длин волн возбуждающего излучения, где эти участки не повреждены.

Процедура восстановления заключается в следующем.

1. Определяется участок, который необходимо заменить. В нашем случае он заранее известен — это пики излучения накачки и рассеяния. В общем случае исследователь должен определить его после установления природы появления ошибок в спектрах. Для этого может быть эффективен МГК. Общим для всех артефактов измерения является их возникновение только на единственном спектре в матрице возбуждения-эмиссии. Таким образом, если значимых ГК больше, чем можно обоснованно предполагать, и среди них присутствует такая, вклад которой существенно выше на единственной длине волны возбуждения, чем на любой другой, значит, спектр, полученный с использованием этой длины волны, нуждается в коррекции в области, где ГК принимает экстремальные значения.

2. Для этого участка и его дополнительной окрестности находится ГК, с необходимой точностью описывающие все прочие спектры, кроме восстанавливаемого. Поскольку для других длин волн возбуждения данные на этом участке спектра не повреждены, найденные ГК будут без ошибок, с точностью до шумов экспериментальной установки, описывать и восстанавливаемый спектр.

3. Проводится сшивка путём нахождения методом наименьших квадратов линейной комбинации ГК, наилучшим образом описывающей дополнительную окрестность.

4. На выбранном участке данные заменяются определённой в п. 3 линейной комбинацией.

На рис. 4, *a* представлен результат подобного восстановления данных в спектре полипропилена, полученного с помощью излучения с длиной волны 270 нм. Спектр нуждался в коррекции в диапазоне длин волн 500–550 нм. Для сшивки применялись окрестности 450–500 и 550–600 нм и четыре главные компоненты, вычисленные нелинейным итерационным методом наименьших квадратов для области 450–600 нм по восьми спектрам, полученным с использованием длин волн возбуждающего излучения 210–260 и 280–290 нм. Видно, что результат хорошо согласуется со спектром, полученным с применением дополнительной фильтрации возбуждающего излучения абсорбционным фильтром УФС1, пропускающим основное излучение и существенно ослабляющим излучение накачки. Как уже упоминалось, из-за широкого диапазона изменения длины волны ультрафиолетового излучения подобное решение затруднительно провести для всех длин волн,

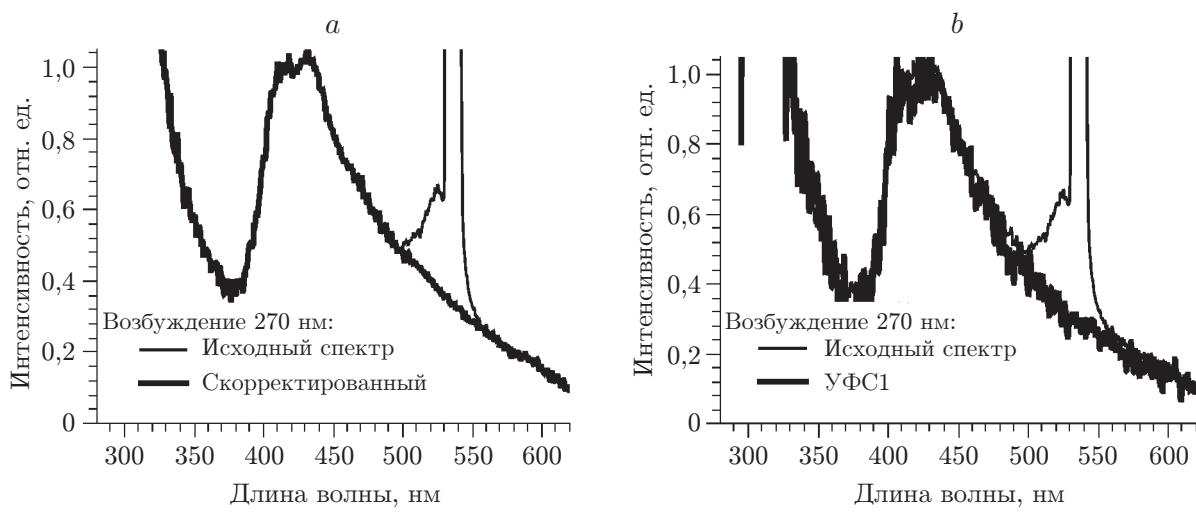


Рис. 4. Сравнение спектров: исходного и восстановленного (a), исходного и спектра с применением дополнительной фильтрации возбуждающего излучения (b)

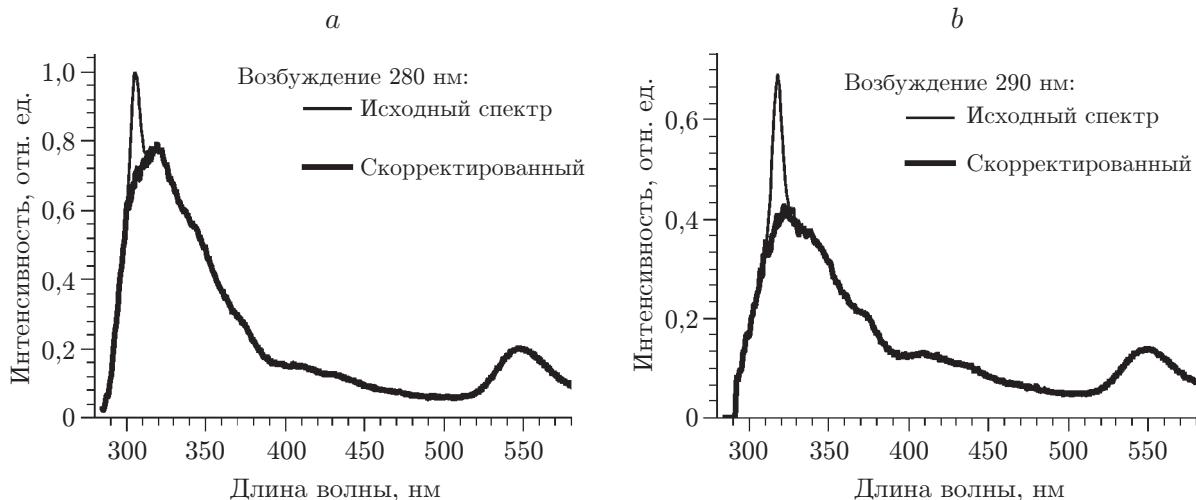


Рис. 5. Сравнение исходного и скорректированных спектров при влиянии рассеяния на регистрируемые спектры при возбуждении 280 (a) и 290 нм (b)

особенно в коротковолновом диапазоне. Причём при использовании фильтра уровень шумов возрастает, поскольку он заметно ослабляет и лазерное излучение.

Кроме того, возможны ситуации, когда от паразитного сигнала нельзя избавиться оптическими методами. В отличие от внешнего излучения накачки (в примере с рассеянием он возникает непосредственно в образце), его пики накладываются на сигнал от флуоресценции. В результате эти пики не могут быть убраны фильтрами и существенно осложняют анализ непосредственно флуоресценции. Тем не менее предлагаемый метод позволяет восстановить повреждённые участки спектра. Так, на рис. 5 представлены восстановленные спектры флуоресценции супензии гомогенизированной ткани аорты коровы при использовании длин волн возбуждения 280 и 290 нм, в которых полностью отсутствуют сигналы рассеяния.

Заключение. Предложен способ очистки данных матриц возбуждения-эмиссии флуоресценции от случайных и систематических артефактов измерения, основанный на методе главных компонент. Способ опирается на такое свойство флуоресценции, как независи-

мость спектров отдельных флуорофоров от длины волны возбуждающего излучения. Благодаря этому он даёт результаты, сравнимые со специальными техническими решениями, однако, отличается универсальностью и не зависит от природы появляющихся артефактов.

Финансирование. Исследование выполнено за счёт гранта Российского научного фонда (проект № 24-22-20078) и при поддержке Правительства Новосибирской области.

СПИСОК ЛИТЕРАТУРЫ

1. Geigle K. P., O'Loughlin W., Hadef R., Meier W. Visualization of soot inception in turbulent pressurized flames by simultaneous measurement of laser-induced fluorescence of polycyclic aromatic hydrocarbons and laser-induced incandescence, and correlation to OH distributions // Appl. Phys. B. 2015. **119**, Iss. 4. P. 717–730.
2. Bhartia R., Hug W. F., Salas E. C. et al. Classification of Organic and Biological Materials with Deep Ultraviolet Excitation // Appl. Spectroscopy. 2008. **62**, Iss. 10. P. 1070–1077.
3. Tsibulskaya E. O., Maslov N. A., Larionov P. M., Ganimedov V. L. Technology of regeneration of the bone tissue in a rotational bioreactor: Fluid flow simulation and laser-induced fluorescence diagnostics // Journ. Appl. Mech. and Tech. Phys. 2020. **61**, Iss. 5. P. 776–786.
4. Maslov N., Larionov P., Rozhin I. et al. Laser-induced fluorescence spectroscopy of the secondary cataract // Optics and Spectroscopy. 2016. **120**, Iss. 6. P. 983–987.
5. Richards-Kortum R., Rava R. P., Cothren R. M. et al. A model for extraction of diagnostic information from laser induced fluorescence spectra of human artery wall // Spectrochimica Acta. Pt. A: Mol. and Biomol. Spectroscopy. 1989. **45**, Iss. 1. P. 87–93.
6. Kirkpatrick N. D., Hoying J. B., Botting S. K. et al. In vitro model for endogenous optical signatures of collagen // Journ. Biomed. Opt. 2006. **11**, Iss. 5. 054021. DOI: 10.1117/1.2360516.
7. Tarai M., Mishra A. K. Application of Multivariate Curve Resolution-Alternate Least Square Technique on Extracting Pure Spectral Components from Multiple Emitting Systems: A Case Study // Journ. Fluorescence. 2017. **27**, Iss. 6. P. 2023–2036.
8. Kumar K. Optimizing Parallel Factor (PARAFAC) Assisted Excitation-Emission Matrix Fluorescence (EEMF) Spectroscopic Analysis of Multifluorophoric Mixtures // Journ. Fluorescence. 2019. **29**, Iss. 3. P. 683–691.
9. Lee Y., Choi S.-I. Data Restoration by Linear Estimation of the Principal Components From Lossy Data // IEEE Access. 2020. **8**. P. 172244–172251.
10. Geladi P., Kowalski B. R. Partial least-squares regression: A tutorial // Analytica Chimica Acta. 1986. **185**. P. 1–17.
11. Tsibulskaya E., Maslov N. Decomposition of multi-component fluorescence spectra by narrow peak method based on principal component analysis // Journ. Chemometrics. 2021. **35**, Iss. 6. DOI: 10.1002/cem.3343.

Поступила в редакцию 28.11.2024

После доработки 16.01.2025

Принята к публикации 17.03.2025