

Алгоритмы коррекции неравномерности освещения в изображениях системы параллельного секвенирования ДНК

В.В. Манойлов¹, А.Г. Бородинов¹, А.И. Петров¹, И.В. Заруцкий¹, А.С. Сараев¹, В.Е. Курочкин¹
¹ ИАП РАН – Институт аналитического приборостроения РАН,
198095, Россия, г. Санкт-Петербург, ул. Ивана Черных, д. 31–33 лит. А

Аннотация

Неравномерное освещение, возникающее на изображениях, полученных во время проведения экспериментов по анализу нуклеиновых кислот в реакционной ячейке системы параллельного секвенирования «Нанофор СПС», приводит к появлению смещения яркости, которое искажает интенсивность и затрудняет количественную оценку флуоресцентных сигналов. Неравномерное освещение может привести к тому, что интенсивность объекта в одной области поля зрения будет измеряться иначе, чем интенсивность объекта с равной концентрацией флуорофора в другой области поля зрения. В работе описываются методы коррекции аддитивной и мультипликативной составляющих факторов, которые искажают регистрируемые изображения. Приводятся результаты исследований по оценке параметров алгоритмов корректировки. В качестве критерия сравнения различных параметров алгоритмов коррекции неравномерного освещения используется количество фрагментов нуклеиновых кислот, которое выбирается программой, осуществляющей сборку целого генома из отдельных его частей. Чем больше таких фрагментов, тем меньше ошибок в результатах построения генома.

Ключевые слова: коррекция фона, секвенирование, обработка изображений, информационные оптические технологии.

Цитирование: Манойлов, В.В. Алгоритмы коррекции неравномерности освещения в изображениях системы параллельного секвенирования ДНК / В.В. Манойлов, А.Г. Бородинов, А.И. Петров, И.В. Заруцкий, А.С. Сараев, В.Е. Курочкин // Компьютерная оптика. – 2026. – Т. 50, № 1. – 1691 – DOI: 10.18287/COJ1691.

Citation: Manoilov VV, Borodinov AG, Petrov AI, Zarutsky IV, Saraev AS, Kurochkin VE. Algorithms for correcting uneven lighting in images of parallel DNA sequencing systems. Computer Optics 2026; 50(1): 1691. DOI: 10.18287/COJ1691.

Введение

Неравномерное освещение, возникающее на изображениях, полученных во время проведения экспериментов в оптической микроскопии и, в частности, по анализу нуклеиновых кислот в реакционной ячейке системы параллельного секвенирования, приводит к появлению смещения, которое искажает интенсивность и затрудняет количественную оценку флуоресцентных сигналов [1, 2, 3, 4].

Существуют различные публикации, в которых обсуждаются причины и отдельные составляющие ошибок, вызванных неравномерностью освещения при измерениях сигналов флуоресценции [1, 2, 3]. В работе [1] показано, что флуоресцентное излучение обычно пропорционально интенсивности освещающего света (за исключением случаев, когда происходит истощение основного состояния флуорофора). Поэтому, если однородный флуоресцентный образец освещается неравномерно, результирующая флуоресценция обычно также будет неравномерной. Неравномерное освещение может быть крайне вредным для количественных измерений, поскольку оно может привести к тому, что интенсивность объекта в одной области поля зрения будет измеряться иначе, чем интенсивность объекта с равной концентрацией флуорофора в другой области поля зрения. Из-за присущих процессу формирования изображения несовершенств микроскопические изображения часто искажаются изменениями интенсивности, проявляющимися в виде больших площадей градиентов интенсивности, отсутствующих в неискаженных изображениях [5]. Это явление еще называют затенением или неоднородностью интенсивности [6], а также называют виньетированием, т. е. ослаблением яркости вдали от центра изображения, что часто приводит к тому, что внешние края изображения оказываются значительно темнее, чем центр [7, 8]. Неравномерное освещение может быть вызвано неточной установкой реакционной ячейки, неравномерность фонового освещения, неравномерностью пространственной чувствительности видеокамеры или ее нелинейности [6]. Согласно работе [8] несколько механизмов могут быть ответственны за эффекты виньетирования. Некоторые из них возникают из-за оптических свойств объективов камеры, наиболее известным из которых является закон падения освещенности по оси. Этот вклад в виньетирование усиливается при съемке под большими углами отклонения от оптической оси [1]. Другие источники виньетирования имеют геометрическую природу. Например, свет, падающий под косыми углами к оптической оси, может быть частично заблокирован диафрагмой или оправой линзы [8]. Виньетирование может быть обусловлено и другими факторами, от освещения, конструкции светового пути между камерой и микроскопом или поведения устройства формирования изображений [1, 8]. Компенсация затенения изображения осуществляется как аппаратными, так и программными методами [10].

В настоящей работе приведены примеры искажений изображений из-за неравномерности освещения в системе параллельного секвенирования «Нанофор СПС», рассмотрены алгоритмы программной коррекции этих искажений с целью повышения точности измерения сигналов флуоресценции. В работе приведены параметры алгоритмов, дающие наибольшее количество пригодных для правильной сборки фрагментов исследуемых нуклеиновых кислот.

Данная публикация является развитием работ по разработкам алгоритмов первичной обработки изображений сигналов флуоресценции для программного обеспечения системы параллельного секвенирования «Нанофор СПС» [11 – 12]. В этих работах и в ссылках в них дается описание принципа действия системы параллельного секвенирования. Для уточнения терминологии, применяемой в следующих подпараграфах данной статьи, приведем определение нескольких терминов и дадим краткое описание эксперимента, проводимого в секвенаторе. Реакционная ячейка, входящая в состав секвенатора «Нанофор СПС», представляет собой микрофлюидное одноканальное устройство, в которое вводится исследуемое вещество и химические реагенты, участвующие в химических реакциях, происходящих в процессе эксперимента. В секвенаторе «Нанофор СПС» имеются четыре видеоканала: А, С, G, Т, по числу нуклеотидов, входящих в состав ДНК. Под действием лазерного излучения возбуждается флуорофор, длина волны которого соответствует определенной букве. В процессе секвенирования генома, содержащего несколько десятков миллионов пар нуклеотидов, снимается серия из нескольких десятков изображений по всей длине реакционной ячейки. Каждое такое изображение называется полем зрения. Съемка изображений на всей длине реакционной ячейки называется циклом или сканом. В нижеследующем тексте номер цикла обозначен буквой c . После завершения текущего цикла производится переход к следующей стадии. В этой стадии через реакционную ячейку пропускают реагенты, отщепляющие флуорофор и останавливающие процесс синтеза. Затем в реакционную ячейку подаются другие реагенты, с помощью которых начинается новый процесс синтеза – новый цикл ($c = c + 1$), где k каждому иммобилизованному кластеру нуклеиновых кислот добавляются новые нуклеотиды. Кластер – группа идентичных ДНК-молекул, прикрепленных к поверхности реакционной ячейки и амплифицированных в одном месте. В секвенаторе «Нанофор СПС» проводится от нескольких десятков до 250 циклов в зависимости от анализируемой пробы и поставленных задач эксперимента.

1. Иллюстрации неравномерного освещения в секвенаторе «Нанофор СПС»

На рис. 1 показано изображение поля зрения из эксперимента на секвенаторе «Нанофор СПС». Центральная область изображения более яркая, а края более темные. Исходный размер изображения имел 1950 строк и 2000 столбцов.

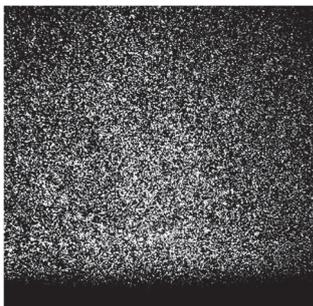


Рис. 1. Типичное изображение поля зрения, демонстрирующее неравномерное освещение

На рис. 2а показан результат обработки изображения, представленного на рис. 1. Изображение, показанное на рис. 1, было обработано с помощью морфологической операции эрозии со структурным элементом *disk* с радиусом 20 пикселей, а затем сглажено с помощью фильтра Гаусса со средним квадратичным отклонением (сигма), равным 55 пикселей. На рис. 2б показано трехмерное изображение после обработки исходного изображения с помощью следующих процедур. Морфологическая обработка и сглаживание с помощью фильтра нижних частот позволили выделить из исходного изображения фоновую составляющую.

На рис. 3а линиями показаны строки изображения, профили которых показаны на рис. 3б. На этом рисунке показаны профили строк с номерами: 100, 1000 и 1850. Как видно из рис. 3б, профили строк изображения имеют разную форму и амплитуду, что подтверждает результат неравномерного освещения. Профиль самой яркой линии соответствует строке 1000 и находится вверху рисунка. Профиль самой тусклой линии соответствует строке 1850 и находится внизу рисунка.

2. Методы коррекции неравномерности освещения

Существует несколько различных подходов к коррекции неравномерного освещения. Методы коррекции могут быть *перспективными*, когда изображение фоновой составляющей, необходимое для калибровки, получено заранее, до проведения эксперимента, или *ретроспективными*, когда изображение фоновой составляющей, необходимое для калибровки, получается в ходе эксперимента, в режиме реального времени [13]. В настоящей работе прежде всего приводятся результаты исследования возможностей ретроспективных методов, так они могут

быть использованы непосредственно в процессе выполнения эксперимента в режиме реального времени. О перспективных методах коррекции неравномерности освещения в изображениях, получаемых в ходе секвенирования на приборе «Нанофор СПС», будет сказано в конце настоящей работы.

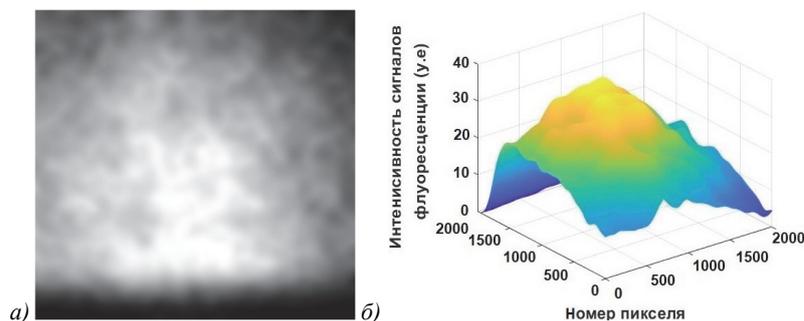


Рис. 2. (а) Результат обработки изображения с помощью морфологической операции эрозии и сглаживанием гауссовым фильтром, (б) трехмерное изображение результата обработки

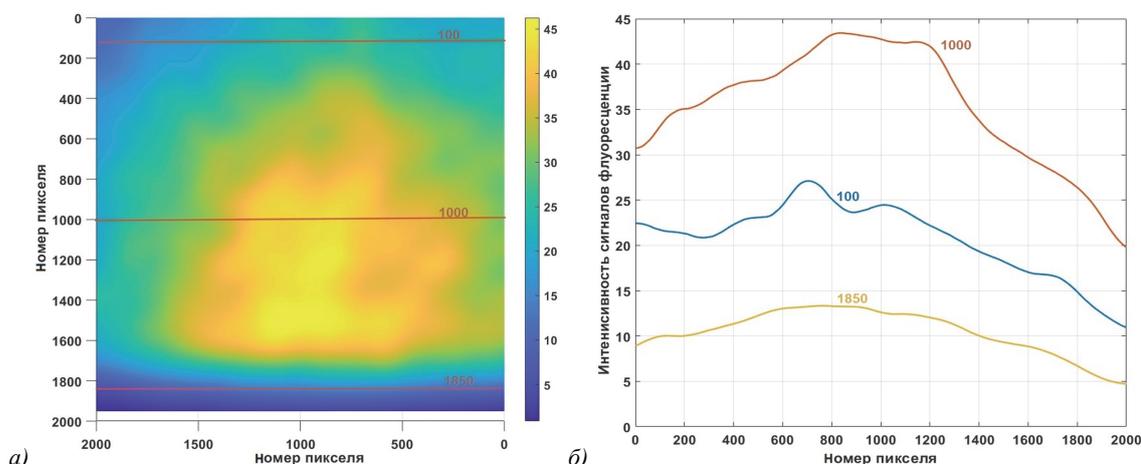


Рис. 3. (а) Результат обработки изображения, представленного на рис. 1, (б) профили обработанных сигналов

Существуют различные ретроспективные методы. Большинство существующих методов коррекции предполагают, что поля смещения в яркостях пикселей изображений являются мультипликативными и медленно меняющимися [14].

Первый класс алгоритмов коррекции применяет фильтрацию с помощью низкочастотных или морфологических операторов, поскольку это простой и интуитивно понятный способ удаления низкочастотных компонентов затенения [13].

Второй класс алгоритмов использует методы поверхностного выравнивания [15], требующие выбора ряда точек на фоне вручную или автоматически, а фон получается путем выравнивания параметрической поверхности [14]. Полиномиальный метод основан на предположении, что изменение интенсивности фонового изображения может быть получено путем подбора полиномиальной функции к значениям интенсивности ряда точек, выбранных на фоне изображения [1]. Он аппроксимирует изображение многочленом и использует отношение ортогональности многочленов Лежандра для расширения изображения как двойной суммы этих функций. Затем сумма оценивается для получения изображения, которое аппроксимирует проекцию на пространство многочленных изображений [1].

Третий класс алгоритмов выполняет минимизацию энтропии [3], поскольку предполагается, что затенение вносит дополнительную информацию в изображение, которая проявляется как более высокая энтропия. Например, в [3] предполагается, что параметрическая полиномиальная поверхность, которая минимизирует энтропию, является компонентом затенения.

3. Линейная модель формирования изображения

Распространенная линейная модель формирования изображения [1, 3, 4], описывающая связь между истинным изображением $U(x; y)$ и полученным изображением $N(x; y)$, выглядит следующим образом:

$$N(x; y) = U(x; y)S_M(x; y) + S_A(x; y), \tag{1}$$

$N(x; y)$ – полученное изображение, искаженное неравномерностью освещения;

$U(x; y)$ – истинное изображение;

$S_M(x; y)$ – мультипликативный компонент затенения;

$S_A(x; y)$ – аддитивный компонент затенения.

Коррекция затенения связана с нахождением исправленного изображения $\hat{U}(x, y)$, которое восстанавливает истинное изображение $U(x, y)$ из полученного изображения $N(x, y)$ [3]

$$N(x, y) \xrightarrow{\text{shading correction}} \hat{U}(x, y) \approx U(x, y). \quad (2)$$

Изображение с коррекцией затенения $\hat{U}(x, y)$ можно рассчитать, используя формулу (1) модели формирования изображения:

$$\hat{U}(x, y) = \frac{N(x, y) - \hat{S}_A(x, y)}{\hat{S}_M(x, y)}, \quad (3)$$

где $\hat{S}_A(x, y)$ и $\hat{S}_M(x, y)$ – оценки аддитивной и мультипликативной компоненты затенения. Таким образом, проблему коррекции затенения можно рассматривать как проблему оценки аддитивной и мультипликативной компоненты затенения [3].

4. Оценка аддитивной компоненты затенения

Для ретроспективной оценки соответствующего компонента аддитивного затенения $\hat{S}_A(x, y)$ исходное изображение обрабатывалось скользящим окном в виде квадрата сначала размером 11×11 и локальный фон был определен как среднее значение самых тусклых 5% пикселей этого пиксельного окна. Затем полученное среднее значение присваивалось всем пикселям окна. В результате для каждого поля зрения формировалось изображение фона, которое вычиталось из исходного изображения. В следующих анализах размер квадрата пиксельного окна тщательно изучался и его размеры составляли следующие значения: 13×13 , 15×15 , 17×17 , 21×21 и 23×23 . В качестве критерия сравнения различных размеров пиксельного окна для выявления наилучшей оценки компонента аддитивного затенения $\hat{S}_A(x, y)$ использовалось количество фрагментов нуклеиновых кислот (картированных ридов), которое было выбрано программой, осуществляющей сборку целого генома из отдельных его частей. Чем больше таких фрагментов, тем меньше ошибок в результатах построения генома. Перед началом такого сравнения были обработаны исходные файлы для 60 циклов в эксперименте по анализу последовательности нуклеотидов в стандартном геноме бактериофага Phix174. В исследуемые исходные файлы сначала не вносилась поправка на неравномерное освещение. Количество фрагментов нуклеиновых кислот, которое было выбрано программой, осуществляющей сборку целого генома из отдельных его частей, оказалось равным 147075. После этого осуществлялась оценка аддитивного компонента затенения и его вычитание из исходного, необработанного изображения.

На рис. 4 приведены результаты выявления количества фрагментов нуклеиновых кислот (картированных ридов), которое было выбрано программой, осуществляющей сборку целого генома из отдельных его частей исследования в зависимости от размеров пиксельного окна для оценки компонента аддитивного затенения $\hat{S}_A(x, y)$, а также от применяемого алгоритма. Кроме алгоритма со скользящим окном, применялись морфологические алгоритмы, реализующие операции эрозии и дилатации со структурными элементами *rectangle* и *disk* разных размеров: 11×11 , 13×13 , 15×15 , 17×17 , 19×19 , 21×21 и 23×23 .

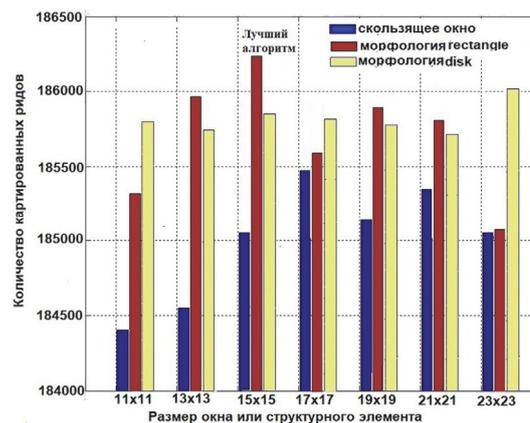


Рис. 4. Влияние размеров пиксельного окна для оценки компонента аддитивного затенения $\hat{S}_A(x, y)$ на количество картированных ридов

На рисунке показано, что наибольшее количество картированных ридов при коррекции аддитивной составляющей затенения имеет алгоритм, реализующий морфологическую обработку со структурным элементом *rectangle* размером 15×15 . Для этого алгоритма количество картированных ридов оказалось равным 186249. Для алгоритма со скользящим окном наилучший результат оказался для окна 17×17 , который составил 185477. Анализ чисел, стоящих на горизонтальной оси рисунка, показывает, что все рассмотренные алгоритмы дают близкие результаты, отличающиеся друг от друга примерно 1%.

Таким образом, введение аддитивного компонента в коррекцию исходных файлов позволило увеличить количество картированных ридов до 186249, что больше чем на 25 % по сравнению с количеством картированных ридов для изображений, которые не подверглись коррекции (147075).

Морфологическая обработка, использующая структурный элемент *disk*, дает близкое к наибольшему количеству картированных ридов, но при этом количество картированных ридов практически не зависит от размера окна.

5. Оценка мультипликативной составляющей затемнения

Для ретроспективной оценки мультипликативного компонента затемнения $\hat{S}_M(x, y)$ могут применяться различные методы, такие как фильтрация Гаусса [1, 16], гомоморфная фильтрация [17, 18] и другие.

Для обработки изображений секвенатора «Нанофор СПС» использовались три метода фильтрации нижних частот, основанных на предположении, что неравномерное освещение является сигналом низкой частоты. Поэтому фильтр нижних частот можно использовать для его извлечения из изображения. Интересующие объекты, т. е. сигналы флуоресценции кластеров ДНК, больше, чем вариация фона, а фон имеет другую интенсивность нежели флуоресценция кластеров. Размытие изображения с помощью фильтра нижних частот основано на том, что все измерения интенсивности являются смесью сигнала и фона. Полученное сглаженное изображение считается оценкой фона изображения [16]. Такая фильтрация может быть выполнена путем свертки исходного изображения $N(x, y)$ с ядром фильтра F_m . В частности, для фильтра Гаусса такое ядро $G(x, y)$ определяется как:

$$G_\sigma = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right), \quad (4)$$

где параметр σ определяет степень размытия изображения. Кроме матрицы с ядром Гаусса, использовались матрицы *average* и *disk* разных размеров. С помощью матрицы типа *average* создается усредняющий фильтр нижних частот. Все компоненты этой матрицы – усредняющие коэффициенты являются одинаковыми числами, значения которых зависят от размера матрицы. С помощью матрицы типа *disk* создается круговой усредняющий фильтр нижних частот внутри квадратной матрицы размером равным удвоенному радиусу. В отличие матрицы типа *average* в матрице *disk* усредняющие коэффициенты находятся внутри круга диаметром равным размеру матрицы, а остальные компоненты этой матрицы равны нулю. Эффект свертки с фильтрующей матрицей заключается в сглаживании значений интенсивности. Цель состоит в том, чтобы сгладить изображение до тех пор, пока не будут сглажены исследуемые объекты флуоресценции. При этом полученное сглаженное изображение будет соответствовать фону.

6. Влияние различных фильтрующих матриц на количество картированных ридов

Оценка $\hat{S}_M(x, y)$ мультипликативного компонента затемнения осуществлялась путем выполнения операции свертки выражения $N(x, y) - \hat{S}_A(x, y)$ с ядром фильтра F_m :

$$\hat{S}_M(x, y) = (N(x, y) - \hat{S}_A(x, y)) * F_m, \quad (5)$$

где знак * – оператор вычисления свертки.

Оценка откорректированного от затемнения изображения $\hat{U}(x, y, c)$ для каждого цикла вычислялась по формуле (6), но с умножением на масштабирующий коэффициент $W(c)$, зависящий от номера цикла c . Коэффициент $W(c)$ вычислялся как среднее значение из максимальных интенсивностей в мультипликативных компонентах затемнения $\hat{S}_M(x, y, c)$ для каждого цикла. Таким образом, оценка изображения $\hat{U}(x, y, c)$, откорректированного от затемнения для каждого цикла c , вычислялась по формуле:

$$\hat{U}(x, y, c) = \frac{N(x, y, c) - \hat{S}_A(x, y, c)}{\hat{S}_M(x, y, c)} W(c). \quad (6)$$

Параметр c введен во все компоненты формулы (6) для того, чтобы показать, что эти компоненты зависят от номера цикла.

На рис. 5 приведены результаты выявления количества фрагментов нуклеиновых кислот (картированных ридов), которое было выбрано программой, осуществляющей сборку целого генома из отдельных его частей исследования в зависимости от типа фильтрующей матрицы, ее размера и значений σ для фильтрации Гаусса. Эти результаты отражают коррекцию от затемнения исходного изображения с учетом адаптивного и мультипликативного компонентов затемнения: $\hat{S}_A(x, y)$ и $\hat{S}_M(x, y)$.

На рисунке показано, что наибольшее количество картированных ридов при коррекции аддитивной и мультипликативной составляющих затемнения имеет алгоритм, реализующий фильтрацию Гаусса с $\sigma = 65$ и размером фильтрующей матрицы 125×125 . При этом размер пиксельного окна для коррекции аддитивной составляющей затемнения составлял 15×15 при использовании в морфологической обработке структурного элемента *rectangle*.

Для этих алгоритмов количество картированных ридов оказалось равным 196098. Что на 5 % выше, чем после введения только аддитивного компонента коррекции затемнения и на 33 % больше, чем с количеством картированных ридов для изображений, которые не подверглись коррекции (147075).

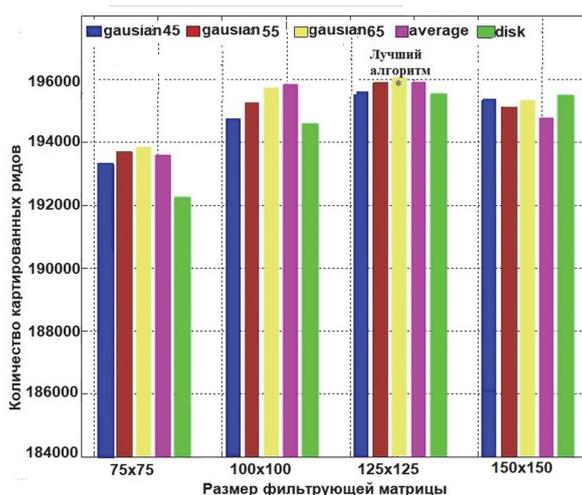


Рис. 5. Количество картированных рядов в зависимости от типа фильтрующей матрицы, ее размера и значений σ для фильтрации Гаусса при коррекции от затемнения изображения с учетом адаптивного и мультипликативного компонентов затемнения $\hat{S}_A(x, y)$ и $\hat{S}_M(x, y)$

7. Результаты применения алгоритмов

Рассмотрим результаты применения алгоритмов, описанных в параграфах 4–6, для обработки данных эксперимента.

На рис. 6 показаны графики профилей 1000-й строки изображения поля зрения С эксперимента на секвенаторе «Нанофор СПС» для исходного файла (верхний график) и для файла после корректировки неравномерности освещения. На рисунке показано, что профиль строки в исходном файле находится на изогнутой вверх линии, тогда как нижний график находится очень близко к «нулевой» линии.

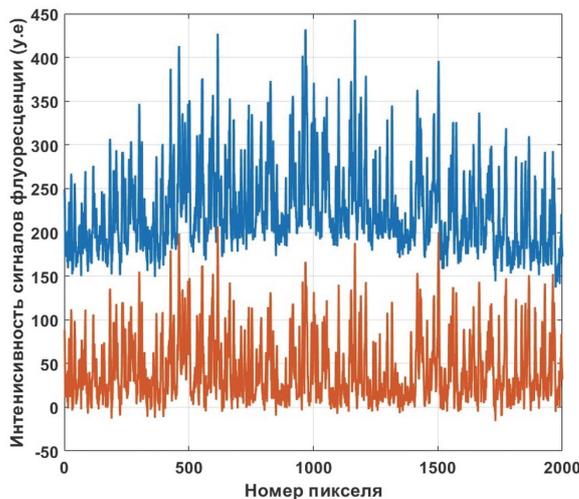


Рис. 6. Графики профилей 1000 строки изображения для исходного файла (верхний график) и для файла после корректировки неравномерности освещения (нижний график) по горизонтальной оси номер пикселя. По вертикальной оси интенсивность флуоресценции в условных единицах

На рис. 7 показаны профили строк с номерами: 100, 1000 и 1850 для откорректированного изображения после того, как к нему были применены те же процедуры обработки, аналогичные показанным на рис. 3б. Профили строк изображения близки к прямым линиям. Средние значения профилей строк 10, 1000 и 1850 близки к нулю и равны соответственно 0,005; 0,001 и 0,03, а их средние квадратичные отклонения равны 0,01; 0,004 и 0,04. Профиль самой тусклой линии соответствует строке 1850, находится внизу рисунка и немного отличается от прямой линии.

На рис. 8 показано трехмерное представление результата обработки изображения, откорректированного на влияние неравномерного освещения. Трехмерное представление имеет форму, близкую к плоскости, и ее среднее значение близко к нулю и равно 0,04. Это обстоятельство свидетельствует о том, что в откорректированном на влияние неравномерного освещения изображении удален фон.

Для сравнения эффективности работы программы, использующей рассмотренные выше алгоритмы коррекции, изображения, полученные в экспериментах на секвенаторе, были обработаны программой, которая реализует алгоритм VaSiC, описанный в работе [4].

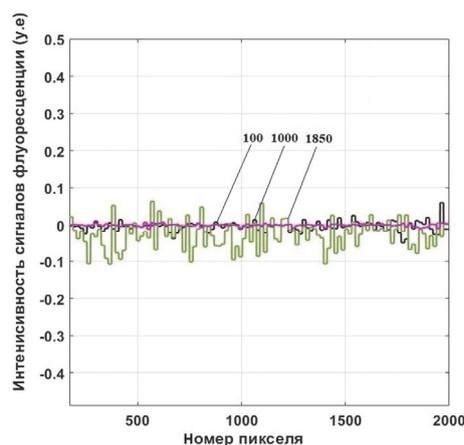


Рис. 7. Профили строк с номерами 100, 1000 и 1850 для откорректированного изображения после применения процедур обработки

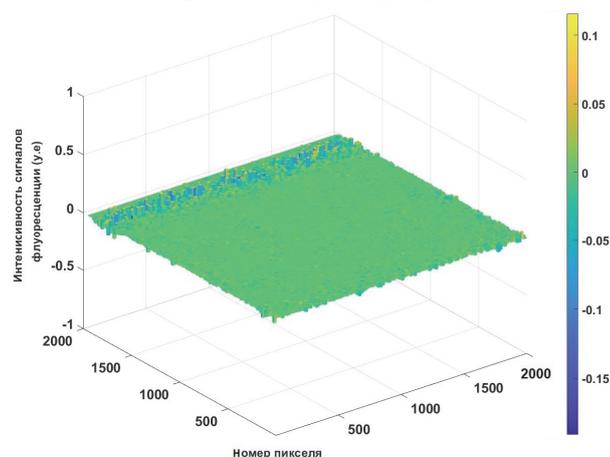


Рис. 8. Трехмерное представление фона, выделенного из откорректированного на влияние неравномерного освещения изображения

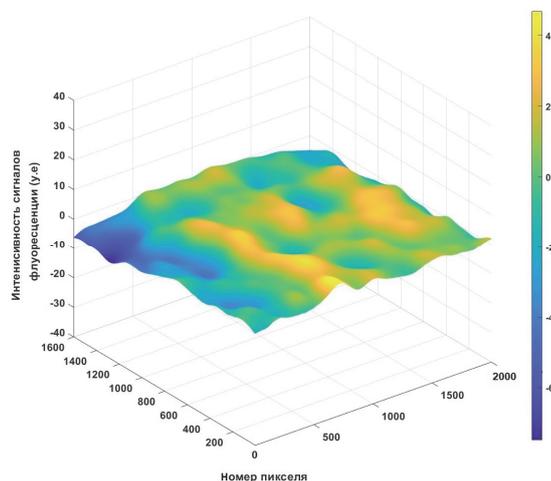


Рис. 9. Трехмерное представление фона, выделенного из откорректированного на влияние неравномерного освещения изображения по программе, реализующей алгоритм BaSiC

Алгоритм BaSiC – это метод коррекции неравномерного освещения и фоновой затенения на изображениях оптической микроскопии. Он предназначен для оценки и коррекции фоновой флуоресценции и неравномерного освещения. BaSiC оценивает два компонента математической модели коррекции освещения: мультипликативный для вариаций интенсивности света и аддитивный для фоновых артефактов. Надо отметить, что для этого алгоритма и для алгоритмов, рассмотренных выше, не требуются эталонные изображения с равномерным освещением.

По программе, реализующей алгоритм BaSiC, были обработаны те же данные эксперимента по построению генома, что и данные для проверки рассмотренных выше алгоритмов. В результате этой обработки было полу-

чено 185083 картированных ридов. При использовании рассмотренных выше алгоритмов количество картированных ридов оказывается примерно на 5 % больше указанной величины, что важно для прецизионного генетического анализа. Одна из причин меньшего количества картированных ридов следующая: фон, выделенный из изображений, откорректированных по алгоритму BaSiC, близок к плоскости, но среднее квадратичное отклонение шума на его поверхности больше чем в 2 раза выше, чем среднее квадратичное отклонение шума на поверхности плоскости фона, выделенной по алгоритмам, рассмотренным выше. Большее значение шума на плоскости фона подтверждается сравнением его трехмерного изображения, представленного на рис. 9, с трехмерным изображением фона, представленным на рис. 8. Другая причина получения меньшего количества картированных ридов заключается в том, что некоторые объекты флуоресценции после коррекции неравномерности освещения оказались смещенными на 1–2 пикселя, что привело к появлению дополнительных ошибок в построении ридов.

8. Перспективные методы коррекции неравномерного освещения в секвенаторе «Нанофор СПС»

При использовании перспективных методов коррекции перед началом эксперимента необходимо заранее получить изображение фоновой составляющей. В ходе выполнения настоящей работы такие априорные изображения фоновой поверхности были получены. Способ получения заключался в том, что в реакционную ячейку, в которую обычно вводится рабочая проба, было введено специальное люминесцирующее вещество – радомин. Время экспозиции было подобрано таким образом, чтобы сигнал флуоресценции мог быть зарегистрирован в каждом канале. Полученные изображения сглаживались с помощью фильтра Гаусса. Затем проводилось суммирование изображений каждого поля зрения и находилось среднее изображение. Такое среднее изображение являлось априорным изображением фоновой поверхности.

Априорные изображения, полученные таким способом, могут быть использованы для сокращения времени выполнения рабочей программы для оценки интенсивностей сигналов флуоресценции в реальном масштабе времени, т.к. для оценки значения аддитивной и мультипликативной составляющих фона не потребуется их вычисления на каждом цикле, а могут быть взяты из данных, предварительно записанных на диск. Однако, как будет показано ниже, мы выигрываем время, но уменьшаем количество правильно картированных ридов, то есть снижаем качество обработки.

Оценка откорректированного на влияние неравномерного освещения для перспективных методов вычислялась по формуле (7), аналогичной ретроспективным методам, только в качестве мультипликативной компоненты затемнения $\hat{S}_M(x, y)$ использовались значения интенсивностей из априорных изображений фоновых поверхностей для каждого канала, а в качестве весового коэффициента W использовались либо средние значения интенсивностей априорных изображений фоновых поверхностей для каждого канала, либо интенсивности, близкие к максимальным (96-процентный перцентиль). Такая операция использовалась для исключения аномальных результатов измерения флуоресцентных сигналов априорных изображений фоновых интенсивностей каждого из каналов (A, C, G, T).

Таким образом, формула для оценки интенсивностей, откорректированных на влияние неравномерного освещения, модифицированная для перспективных методов, коррективки, имеет вид:

$$\hat{U}(x, y, c) = \frac{N(x, y, c) - \hat{S}_A(x, y, c)}{\hat{S}_M(x, y)} W. \quad (7)$$

В этой формуле, в отличие от формулы (6) мультипликативный компонент затемнения $\hat{S}_M(x, y)$ и весовой коэффициент W не зависят от номера цикла C .

Проверка возможностей использования априорных изображений фоновой составляющей для коррекции неравномерностей освещения показала, что количество фрагментов нуклеиновых кислот (картированных ридов) оказалось равным 178744. Что примерно на 10 % меньше, чем при коррекции ретроспективными методами, описанными выше.

Выводы

1. Использование перспективных и ретроспективных методов коррекции неравномерности освещения в изображениях системы параллельного секвенирования «Нанофор СПС» позволило увеличить количество фрагментов нуклеиновых кислот (картированных ридов), которое было выбрано программой, осуществляющей сборку целого генома из отдельных его частей более чем на 25 %, а в случае использования ретроспективных методов с наилучшими параметрами – до 33 %.

2. При оценке аддитивного компонента в алгоритме коррекции неравномерности освещения в изображениях секвенатора выявлено, что наилучшие результаты дает алгоритм с использованием морфологической обработки со структурным элементом *average* размером 15×15 пикселей для построения изображения фона.

3. При оценке мультипликативного компонента в алгоритме коррекции неравномерности освещения в изображениях секвенатора выявлено, что наилучшие результаты дает алгоритм фильтрации нижних частот на основе функции Гаусса со средним квадратичным отклонением 65 пикселей и размером фильтрующей матрицы 125×125 пикселей.

4. Ретроспективные методы коррекции дают примерно на 10 % лучший результат по коррекции изображений, искаженных неравномерностью освещения по сравнению с перспективными методами.

Благодарности

Работа выполнена при поддержке Министерства науки и высшего образования в рамках выполнения работ по Государственному заданию № 075–00444–25–00 (от 26.12.2024).

References

- [1] Wolowski VR. High-quality, high-throughput measurement of protein–DNA binding using HiTS–FLIP. Dissertation zur Erlangung des Doktorgrades der Fakultät für Chemie und Pharmazie der Ludwig–Maximilians–Universität München; 2016. 251. 10.5282/edoc.19445
- [2] Waters JC. Accuracy and precision in quantitative fluorescence microscopy. *The Journal of cell of biology* 2009; 185(7): 1135–1148.
- [3] Likar B, Pernu F. Retrospective shading correction of microscopical images. Retrospective shading correction of microscopical images. In Proc. Czech Pattern Recognition Workshop, Eds. Svoboda, PT; 2000: 15–20.
- [4] Peng T, Thorn K, Schroeder T, Wang L, Theis FJ, Marr C, Navab N. A BaSiC tool for background and shading correction of optical microscopy images. 2017. NATURE COMMUNICATIONS. DOI: 10.1038/ncomms14836
- [5] Inoue S. Video microscopy. Springer Science & Business Media; 1986: 460.
- [6] Likar B, Pernu F. Retrospective shading correction of microscopical images. Retrospective shading correction of microscopical images. In Proc. Czech Pattern Recognition Workshop, Eds. Svoboda, PT; 2000: 15–20.
- [7] Marty GD. Blank–field correction for achieving a uniform white background in brightfield digital photomicrographs. *BioTechniques* 2007; 42:716–720. DOI 10.2144/000112488
- [8] Zheng Y, Lin S, Kambhamettu C, Yu J, Kang SB. Single–image vignetting correction. *IEEE transactions on pattern analysis and machine intelligence* 2009; 31(12): 2243–2256
- [9] Zheng Y, Lin S, Kambhamettu C, Yu J, Kang SB. Single–image vignetting correction. *IEEE transactions on pattern analysis and machine intelligence* 2009; 31(12): 2243–2256
- [10] Digital Imaging – Microscopy Capabilities [Electronic resource]. URL: https://stormoff.ru/mediacenter/article/article_84 (access data: 18.01.2025)
- [11] Manoilov VV, Borodinov AG, Petrov AI, Zarutskiy IV, Bardin BV, Yamanovskaya AY, Saraev AS, Kurochkin VE. Construction of a nucleotide sequence using machine learning methods in the Nanofor SPS" sequencer. *Technical Physics*, 2024; 69(9): 1436–1444. DOI: 10.61011/TP.2024.09.59293.35–24.
- [12] Borodinov AG, Manoilov VV, Zarutsky IV, Petrov AI, Kurochkin VE. Generations of DNA sequencing methods (review). *Nauchnoe Priborostroyeniye* [Scientific Instrumentation], 2020, vol. 30, iss. 4, pp. 3–20.
- [13] Reyes–Aldasoro CC. Retrospective shading correction algorithm based on signal envelope estimation. *Electronics Letters* 2009; 45(9): 454–456. DOI: 10.1049/el.2009.0320
- [14] Kubecka L, Jan J, Kolar R. Retrospective illumination correction of retinal images. *International journal of biomedical imaging* 2010; Article ID 780262: 10. DOI:10.1155/2010/780262
- [15] Hou Z, Huang S, Hu Q, Nowinski WL. A fast and automatic method to correct intensity inhomogeneity in MR brain images. In *Medical Image Computing and Computer–Assisted Intervention–MICCAI* 2006; 324–331. DOI: 10.1007/11866763_40.
- [16] Leong FJW–M, Brady M, McGee JO. Correction of uneven illumination (vignetting) in digital microscopy images. *Journal of clinical pathology* 2003; 56(8): 619–621. of *clinical pathology* 2003; 56(8): 619–621.
- [17] Etemadnia H, Asharif MR. Homomorphic filtering approach using HSV color space in automatic image shadow segmentation. *WSEAS Trans Systems* 2004; 3: 1150–1154.
- [18] Wen–Cheng W, Xiao–Jun C. A Segmentation Method for Uneven Illumination Particle Images. *Research Journal of Applied Sciences, Engineering and Technology* 2013; 5(4):1284–1289.

Сведения об авторах

Манойлов Владимир Владимирович, д.т.н., доцент, заведующий лабораторией автоматизации измерений и цифровой обработки сигналов ИАП РАН. Область научных интересов: представление и обработка сигналов и изображений в аналитических приборах. Число научных публикаций – 101. E-mail: manoilov-vv@mail.ru

Бородинов Андрей Геннадьевич, к.ф–м.н., старший научный сотрудник в лаборатории методов и приборов иммунного и генетического анализа ИАП РАН. Область научных интересов: математическая статистика, проблемы анализа, обработки и представления данных, искусственный интеллект. Число научных публикаций – 15. E-mail: borodinov@gmail.com

Петров Александр Иванович, к.т.н., заведующий сектором электроники и программного обеспечения в лаборатории методов и приборов иммунного и генетического анализа ИАП РАН. Область научных интересов: представление и обработка сигналов и изображений в аналитических приборах. Число научных публикаций – 21. E-mail: fataip@mail.ru

Заруцкий Игорь Вячеславович, к.т.н., старший научный сотрудник в лаборатории автоматизации измерений и цифровой обработки сигналов ИАП РАН. Область научных интересов: представление и обработка сигналов и изображений в аналитических приборах. Число научных публикаций – 41. E-mail: igorzy@yandex.ru

Сараев Алексей Сергеевич, инженер 1-й категории в лаборатории методов и приборов иммунного и генетического анализа ИАП РАН. Область научных интересов: представление и обработка сигналов и изображений в аналитических устройствах. Количество научных публикаций – 7. E-mail: alex.niispb@yandex.ru

Курочкин Владимир Ефимович, д.т.н., профессор, руководитель научного направления, заведующий лабораторией методов и приборов иммунного и генетического анализа ИАП РАН. Область научных интересов: исследования и оптимизация электромиграционных методов анализа, развитие аналитических методик для капиллярного электрофореза, исследование оптических методов детектирования, разработка методов и приборов для ДНК-анализа, разработка методик подготовки проб и специализированных реактивов. Число научных публикаций – 200. E-mail: lavrovav@yandex.ru

ГРНТИ: 29.03.45

Поступила в редакцию 24 февраля 2025 г. Окончательный вариант – 23 апреля 2025 г.

Algorithms for correcting uneven lighting in images of parallel DNA sequencing systems

V.V. Manoilov¹, A.G. Borodinov¹, A.I. Petrov¹, I.V. Zarutsky¹, A.S. Saraev¹, V.E. Kurochkin¹

¹IAI RAS – Institute for Analytical Instrumentation of the Russian Academy of Sciences,
198095, Russia, St. Petersburg, Ivan Chernykh St. 31–33 lit. A

Abstract

The uneven illumination that occurs in the images obtained during experiments on the analysis of nucleic acids in the flow cell of the parallel sequencing system "Nanofor SPS" leads to the appearance of bias, which distorts the intensity and makes it difficult to quantify the fluorescent signals. Uneven lighting can cause the intensity of an object in one area of the visual field to be measured differently from the intensity of an object with an equal concentration of fluorophore in another area of the visual field. The paper describes methods for correcting the additive and multiplicative components of factors that distort the measured images. The results of research on the evaluation of parameters of correction algorithms are presented. The number of fragments of nucleic acids is used as a criterion for comparing various parameters of algorithms for correcting uneven lighting, which is selected by a program that assembles the whole genome from its individual parts. The more such fragments there are, the fewer errors there are in the results of genome construction.

Keywords: background correction, sequencing, image processing, information optical technologies.

Citation: Manoilov VV, Borodinov AG, Petrov AI, Zarutsky IV, Saraev AS, Kurochkin VE. Algorithms for correcting uneven lighting in images of parallel DNA sequencing systems. *Computer Optics* 2026; 50(1): 1691. DOI: 10.18287/COJ1691.

Acknowledgements: This work was supported by the Ministry of Science and Higher Education within the State assignment № 075–00444–25–00 (by 26.12.2024).

Author's information

Vladimir Manoilov, Dr. Sci., associate professor, head of Laboratory of automation of measurements and digital signal processing of IAI RAS. Research interests: the representation and processing of signals and images in analytical devices. The number of scientific publications – 101. E-mail: manoilov-vv@mail.ru

Andrew Borodinov, Ph.D. in Physics and Mathematics, Senior Researcher in the laboratory of methods and instruments for immune and genetic analysis of the IAI RAS. Research interests: mathematical statistics, problems of analysis, processing and presentation of data, artificial intelligence. The number of scientific publications – 15. E-mail: borodinov@gmail.com

Alexander Petrov, Ph.D., head of sector of electronics and software in the laboratory of methods and instruments for immune and genetic analysis of the IAI RAS. Research interests: the representation and processing of signals and images in analytical devices. The number of scientific publications – 41. E-mail: fataip@mail.ru

Igor Zarutsky, Ph.D., Senior Researcher at laboratory of automation of measurements and digital signal processing of IAI RAS. Research interests: the representation and processing of signals and images in analytical devices. The number of scientific publications – 51. E-mail: igorzv@yandex.ru

Alexey Saraev, engineer of the 1st category in the laboratory of methods and instruments for immune and genetic analysis of the IAI RAS. Research interests: representation and processing of signals and images in analytical devices. The number of scientific publications – 7. E-mail: alex.niispb@yandex.ru

Vladimir Kurochkin, Dr. Sci., professor, head of the scientific direction, head of the laboratory of methods and instruments for immune and genetic analysis of the IAI RAS. Research interests: research and optimization of electromigration analysis methods, the development of analytical methods for capillary electrophoresis, the study of optical methods of detection, the development of methods and instruments for DNA analysis, the development of methods for preparing samples and specialized reagents. The number of scientific publications – 200. E-mail: lavrovvas@yandex.ru

Received February 24, 2025. The final version – April 23, 2025.
