

(19)



**Евразийское
патентное
ведомство**

(21) **201990595** (13) **A1**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2020.07.20

(51) Int. Cl. **G06K 9/62** (2006.01)
G02B 21/10 (2006.01)

(22) Дата подачи заявки
2018.07.02

**(54) СПОСОБ ДЛЯ АВТОМАТИЗИРОВАННОГО НЕКОНТРОЛИРУЕМОГО
ОНТОЛОГИЧЕСКОГО ИССЛЕДОВАНИЯ СТРУКТУРНЫХ ВИДОВ В ЭЛЕКТРОННЫХ
МИКРОФОТОГРАФИЯХ**

(31) 62/535,851

(32) 2017.07.22

(33) US

(86) PCT/US2018/040625

(87) WO 2019/022925 2019.01.31

(71) Заявитель:

**ИНТЕЛЛИДЖЕНТ ВАЙЕРЭС
ИМЭДЖИНГ ИНК. (US)**

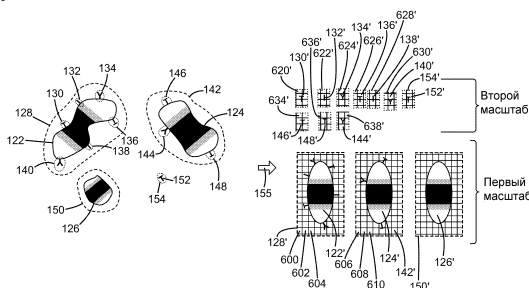
(72) Изобретатель:

Райнер Мартин (SE)

(74) Представитель:

Нагорных И.М. (RU)

(57) Способ заключается в разделении темных объектов (102, 122, 123, 126), подструктур и фона изображения электронного микроскопа (304) на сегменты (128, 142, 150) путем анализа значений пикселей. Сегменты преобразуются и выравниваются таким образом, чтобы преобразованные объекты, подструктуры и фон были достоверно сопоставимы. Преобразованные сегменты (128', 142', 150') сгруппированы в классы, которые используются для онтологического исследования образцов, визуализируемых с помощью электронной микроскопии. Сравнение неравенства треугольников может быть использовано для дальнейшей кластеризации групп объектов, чтобы перенести понимание из различных взаимодействий между объектами и ассоциировать взаимодействия друг с другом.



A1

201990595

201990595

A1

СПОСОБ ДЛЯ АВТОМАТИЗИРОВАННОГО НЕКОНТРОЛИРУЕМОГО ОНТОЛОГИЧЕСКОГО ИССЛЕДОВАНИЯ СТРУКТУРНЫХ ВИДОВ В ЭЛЕКТРОННЫХ МИКРОФОТОГРАФИЯХ

5 Область техники

Изобретение относится к способу автоматизированного неконтролируемого онтологического исследования проявлений структур, таких как органические и неорганические материалы, в электронных микрофотографиях.

10 История вопроса и краткое описание изобретения

В прошлом предпринимались многочисленные усилия по выявлению, классификации и анализу клеточных структур, включающих вирусы и другие компоненты. Для описания, сегментации и классификации вирусов с помощью имеющихся графических технологий были разработаны различные способы анализа изображений. Обычно пользователь заранее решает, что именно он будет искать под микроскопом и при каком увеличении. В прошлом были разработаны различные способы обнаружения, которые адаптированы для идентификации, измерения и т.д. того, что ищет пользователь. При использовании специальных и заранее определенных способов обнаружения можно легко пропустить важную информацию на изображении, которая не обнаружена с помощью специальных способов обнаружения. Кроме того, для обеспечения определенного уровня качества сложно последовательно осуществлять процессы идентификации и анализа вручную. В прошлом было трудно объективно, многократно и достоверно идентифицировать, классифицировать и характеризовать объекты, полученные с помощью микроскопических изображений. Предшествующие процессы предварительной идентификации и классификации вирусных частиц были в значительной степени подвержены влиянию пользователя и трудоемки в использовании, поскольку способы из уровня техники требуют от пользователя принятия решений о том, как сегментировать изображения и какие масштабы/увеличение микроскопического исследования должны быть использованы. Структурированные способы с искусственным интеллектом, использующие фиксированные банки фильтров, такие как современные свёрточные нейронные сети, часто сталкиваются с проблемами естественного выравнивания, что делает разрешение информации более грубым, чем то, что оптимально достижимо. Предшествующие решения

опираются на источники больших данных, чтобы попытаться компенсировать грубость, но, как следствие, они часто оказываются неэффективными и требуют много времени для извлечения новых знаний.

5 Существует потребность в более совершенном и надежном способе исследования таких объектов, как вирусные частицы, который был бы автоматическим, но при этом достаточно гибким для работы с большим разнообразием биологических объектов и образцов.

10 Способ согласно настоящему изобретению обеспечивает решение вышеуказанных задач. Более конкретно, способ предназначен для автоматизированной онтологической группировки объектов. Цифровая камера делает снимок или фотографию изображения в электронном микроскопе. Снимок или фотография сохраняется в компьютере или процессоре в виде цифрового изображения. Можно также сделать много снимков. При первом масштабе увеличения компьютер автоматически идентифицирует объекты на цифровом изображении, сравнивая яркость пикселей на цифровом изображении. Компьютер определяет степень ассоциации между каждым пикселем в зависимости от яркости каждого пикселя. Компьютер формирует первый сегмент вокруг первого идентифицированного объекта и второй сегмент вокруг второго идентифицированного объекта на изображении разделенными, где пиксели имеют наименьшую связь друг с другом. Компьютер математически преобразует первый идентифицированный объект в первый преобразованный объект в первом преобразованном сегменте, а второй идентифицированный объект во второй преобразованный объект во втором преобразованном сегменте, при этом преобразованные первый и второй объекты имеют формы с фиксированной ориентацией. На основе яркости пикселей компьютер определяет первую топологию первого преобразованного объекта. Компьютер сравнивает яркость пикселей в первом преобразованном сегменте с яркостью пикселей во втором преобразованном сегменте.

30 Во втором варианте настоящего изобретения компьютер при втором масштабе увеличения, когда второй масштаб увеличения выше, чем первый масштаб увеличения, автоматически идентифицирует подструктуры на цифровом изображении путем сравнения яркости пикселей на цифровом изображении. Компьютер определяет степень ассоциации между каждым пикселем в зависимости от яркости каждого пикселя. Компьютер формирует первый сегмент подструктуры вокруг первой идентифицированной подструктуры, а второй сегмент

подструктуры вокруг второй идентифицированной подструктуры на цифровом изображении разделенными, где пиксели имеют наименьшую связь друг с другом. Компьютер математически преобразует первую идентифицированную подструктуру в первую преобразованную подструктуру в первом преобразованном сегменте подструктуры, а вторую идентифицированную подструктуру во вторую преобразованную подструктуру во втором преобразованном сегменте подструктуры, при этом преобразованные первая и вторая подструктуры имеют формы с фиксированной ориентацией. На основе яркости пикселей компьютер определяет вторую топологию первой преобразованной подструктуры. Компьютер сравнивает яркость пикселей в первом преобразованном сегменте подструктуры с яркостью пикселей во втором преобразованном сегменте подструктуры.

В альтернативном варианте компьютер определяет первую топологию второго преобразованного объекта и вторую топологию второй преобразованной подструктуры. Основываясь на топологиях, компьютер назначает координаты первому преобразованному объекту и вторым преобразованным объектам таким образом, чтобы первый преобразованный объект и второй преобразованный объект находились в непосредственной близости в перцептивном пространстве и назначает координаты первой преобразованной подструктуре и второй преобразованной подструктуре таким образом, чтобы первая преобразованная подструктура и вторая преобразованная подструктура находились в непосредственной близости в перцептивном пространстве, но на удалении от первого и второго преобразованных объектов.

В альтернативном варианте компьютер идентифицирует третий объект и формирует третий сегмент вокруг третьего объекта. Компьютер определяет, что первый сегмент и первый сегмент подструктуры имеют общий пиксел, и второй сегмент и второй сегмент подструктуры имеют общий пиксел. Компьютер определяет, что третий сегмент не имеет пикселя, общего с подструктурой. Компьютер назначает координаты первого и второго сегментов таким образом, чтобы первый и второй сегменты находились ближе друг к другу в перцептивном пространстве, чем третий сегмент - к первому и второму сегментам.

В другом варианте компьютер идентифицирует объекты при первом масштабе увеличения одновременно с идентификацией подструктур при втором масштабе увеличения.

В другом варианте компьютер идентифицирует третий объект на втором цифровом изображении второй фотографии, отличной от цифрового

изображения.

Краткое описание чертежей

На Фиг. 1 изображены вирусные частицы на электронной микрофотографии;

- 5 На Фиг. 2А показаны сегментированные при первом более мелком масштабе вирусные частицы;

На Фиг. 2В показана Фиг.2А, сегментированная при втором более крупном масштабе;

- 10 На Фиг. 3 схематично показан вид вирусных частиц и подструктур (поверхностных белков);

На Фиг. 4 схематично показан вид вирусных частиц и поверхностных белков, которые преобразованы независимо с фиксированной ориентацией;

На Фиг. 5 схематично показан вид групп преобразованных вирусных частиц и подструктур;

- 15 На Фиг. 6 схематично показан вид дополнительных ассоциаций объектов внутри групп;

На Фиг. 7 схематично показан вид кластеризации внутри групп на основе ассоциации между объектами в группах;

- 20 На Фиг. 8 схематично показан вид электронного микроскопа согласно настоящему изобретению;

На Фиг. 9 схематично показан вид вирусной частицы;

На Фиг. 10 схематично показан вид групповой кластеризации, где измеряются взаимосвязи; и

- 25 На Фиг. 11 показан еще один вид групповой кластеризации, где измеряются взаимосвязи.

Подробное описание

- В целом, настоящее изобретение является способом (1) разделения необработанных данных, в данном случае изображений различных типов органических или неорганических материалов из изображений электронной микроскопии на значимые сегменты, (2) выравнивания сегментов таким образом, чтобы они были достоверно сопоставимы, и (3) разделения сегментов на классы, которые могут быть использованы для онтологического исследования образцов, визуализируемых с помощью электронной микроскопии. Одним из важных аспектов настоящего изобретения является то, что структурная информация на
- 30
- 35

изображении разделяется на связанные подмножества с собственными полями выравнивания, основанными на самих данных, а не на какой-либо фиксированной или заранее определенной структуре или фиксированных фильтрах, что характерно для большинства способов из предшествующего уровня техники.

- 5 На практике цифровая камера 300 (лучше всего показана на Фиг. 8) используется для получения снимка или фотографии изображения 302 (или многих изображений) в электронном микроскопе 304. Камера 300 может, конечно, делать много снимков изображения 302 или изображений. Затем снимок сохраняется в CPU (центральном процессоре) или компьютере 306 в виде
- 10 цифрового изображения. Цифровое изображение 100, показанное на Фиг.1-2В, является наглядным примером цифрового изображения, сохраненного в компьютере 306. Затем компьютер или CPU 306 используется для анализа цифрового изображения 100, проходя через каждый пиксел цифрового изображения 100, чтобы, помимо прочего, определить яркость каждого пиксела.
- 15 Компьютер, предпочтительно, выполняет математическое преобразование объектов на цифровом изображении 100, которые используются для дальнейшего анализа изображения, как это подробно описано ниже.

Способ разделения/сегментации и выравнивания необработанных данных выполняется путем исследования спектра ядра, описанного ниже:

- 20 Исследования спектра ядра

$$g(y, z) = \phi(c_1 d(I(y), I(z))^2 + s(z) * d(x(y), x(z))^2 + c_2 d(s(y), s(z))^2) \quad (1)$$

По интегральному уравнению

$$(x(z), s(z)) = \left(\frac{\int g(y, z) x(y) dy}{N_x}, \frac{\int g(y, z) s(y) dy}{N_s} \right)$$

Где

- 25 (I) – сигнал изображения,
 (x) – собственный вектор сегментации и выравнивания,
 (s) – собственный вектор масштабирования,
 (y) и (z) – пространственный индекс дискретного изображения в ограничивающем квадрате $B = [0, \text{ширина}] \times [0, \text{высота}] \times [0, S_{\max}]$,
- 30 (N_x) – коэффициент нормализации для собственного вектора x,
 (N_s) – коэффициент нормализации для собственного вектора s,
 (d) – стандартная евклидова норма,
 (ϕ) – монотонно возрастающая функция, такая как экспоненциальная функция (e),

(с1) и (с2) – весовые коэффициенты.

Одним из ключевых моментов данного нововведения является обратный ход масштабирования (s) и сегментации (x) в самом ядре для генерации разделения. Две самые большие собственные функции оператора (g) используются для поиска фиксированных точек, т.е. все (y), имеющие одинаковое значение (x).

$$S_c = \{y \in B: X(y) = c\}$$

Каждое непустое множество называется сегментом (S), обозначенным как (с) или просто как (S_c). Множество всех сегментов называется (Σ). Множество Σ является предметом онтологического исследования, описанного ниже.

Одним из первых этапов настоящего изобретения является одновременная идентификация каждого объекта на изображении, включая фон, и разделение изображения на сегменты в разных масштабах. Поскольку размер объекта(ов) неизвестен и может незначительно изменяться в зависимости от подготовки образца, инструментария и эффекта выборки, сначала используются все масштабы увеличения микроскопа, пока не будет определено, какие масштабы являются наиболее подходящими. Отдельные крупные объекты, такие как крупные вирусные частицы, и агрегаты мелких вирусных частиц могут быть лучше всего видны при определенном увеличении. Вирусные частицы обычно находятся в диапазоне размеров от 20 нм до 300-400нм. На этапе сегментации способ может использоваться для автоматического определения наиболее соответствующего или подходящего масштаба в зависимости от размера обнаруженного объекта. Способ настоящего изобретения не зависит от предварительно установленного масштаба увеличения, который, например, устанавливается пользователем. Вместо этого способ анализирует и определяет размеры объектов, анализируя изменение яркости каждого пикселя за пикселем, и на основе этого анализа определяет, где находятся края объектов и какое увеличение является наиболее оптимальным или подходящим, чтобы масштаб изображения соответствовал размеру каждого объекта, найденного в определенном месте на изображении. Предполагается, что пиксели со сходной яркостью (т.е. сильной ассоциацией) внутри краев принадлежат одному и тому же объекту, т.е. частице вируса. Вместо определения сходства на основе яркости можно также использовать другие критерии, такие как сходство в цвете.

Пиксели со сходной яркостью (т.е. сильной ассоциацией) на заднем плане вне объекта также считаются принадлежащими фону. Поскольку каждый пиксел

анализируется на изображении, естественная точка прерывания в конечном итоге обнаруживается там, где существует относительно слабая ассоциация или сходство между пикселями, например, вблизи или на краю объекта. В качестве наглядного примера можно привести пиксел p_1 со значением яркости 5, пиксел p_2 со значением яркости 5.1, пиксел p_3 со значением 10, пиксел p_4 со значением 43, пиксел p_5 со значением 50 и пиксел p_6 со значением 51. Анализ математическим алгоритмом определяет, что для пикселя p_3 существует более сильная ассоциация с пикселем p_2 , чем с пикселем p_4 , поскольку его значение яркости ближе к пикселу p_2 , чем к пикселу p_4 . Для пикселя p_4 существует более сильная ассоциация с пикселем p_5 , чем с пикселем p_3 , поскольку значение яркости пикселя p_4 ближе к значению яркости пикселя p_5 , чем к пикселу p_3 . Именно здесь создается сегментация (между пикселями p_3 и p_4) потому, что пиксели внутри линии сегментации больше ассоциированы или схожи друг с другом, чем с пикселями вне линии сегментации, а пиксели вне линии сегментации также больше ассоциированы друг с другом, чем с пикселями внутри сегментации, поскольку разница в яркости каждого пикселя внутри сегментации меньше, чем снаружи, и наоборот. Например, если бы пиксел p_6 имел яркость 113, а пиксел p_7 имел яркость 115, то пиксел p_5 (яркость 50) был бы больше ассоциирован с пикселем p_4 (яркость 43), чем с пикселем p_6 (яркость 113) и сегментация была бы между пикселем p_5 и пикселем p_6 , то есть пиксел p_5 «последовал» бы за пикселем p_4 и принадлежал бы объекту.

Если изображение имеет последовательность пикселей p_1, p_2, p_3 и т.д., способ сравнивает яркость между пикселями p_1 и p_2, p_2 и p_1, p_2 и p_2, p_2 и p_3, p_3 и p_4 . Он определяет создание линии сегментации между пикселями p_2 и p_3 , когда пиксел p_2 больше похож на пиксел p_1 , а p_3 больше похож на p_4 , чем p_2 похож на p_3 .

Таким образом, следует отметить, что заранее заданное пороговое значение не используется для определения того, где должен находиться край или разделительная сегментация. В способе настоящего изобретения основное внимание уделяется наиболее сильной ассоциации между значениями пикселей, но не там, где существует наибольшая разница в яркости. Другими словами, важной особенностью настоящего изобретения является то, что сегментация не является фиксированной или предопределенной, а точная сегментация зависит от размера объектов на изображении, определяемого ассоциациями между пикселями на изображении. Это означает, что масштаб увеличения, выбранный

как наиболее оптимальный (в зависимости от размера сегментов), может варьироваться для каждого места на изображении, поэтому при анализе изображения используется несколько масштабов одновременно. Определенный масштаб увеличения может быть оптимальным для сегментирования целых объектов или кластеров объектов, в то время как другой масштаб увеличения может быть использован для определения связи между пикселями внутри объекта таким образом, что изображение разделяется на меньшие сегменты для отображения деталей, таких как структура или форма вирусной частицы и белков, связанных с вирусной частицей. Также, как подробно описано ниже, после преобразования объектов на изображении, объекты нормализуются, и объекты с одинаковой топологией будут выглядеть более похоже друг на друга, чтобы сделать ассоциацию между объектами сильнее, несмотря на различия в размере и форме до преобразования. Затем для определения ассоциаций проводится такой же анализ каждого пикселя преобразованных объектов и подструктур в различных масштабах. Это делает способ достаточно нечувствительным к изменчивости размеров и формы анализируемых объектов.

На Фиг. 1 показано цифровое изображение 100 вирусных частиц 102, а также других частиц и мусора на электронной микрофотографии. Цифровое изображение 100 содержит множество пикселей, т.е. мельчайших областей подсветки, из которых состоит изображение 100. На Фиг. 2А показано изображение 100, разделенное на сегменты 104 при первой более крупном масштабе увеличения, а на Фиг. 2В показано изображение, разделенное на сегменты 106 при второй отличном и более мелком масштабе увеличения. Другими словами, изображение одновременно делится на сегменты, на основе локальных различий при разных масштабах увеличения. Разделительные линии между сегментами 104, 106 нарисованы там, где есть разрыв или изменение изображения, например, края вокруг черного непрерывного объекта (например, частицы вируса 102) на сером фоне 110. Как упоминалось ранее, важной особенностью настоящего изобретения является то, что для создания сегментов нет необходимости использовать какие-либо пороговые значения. Другим важным принципом настоящего изобретения является упор на ассоциации/подобия, а не на сходства или различия между объектами в сегментах.

Поскольку используются все возможные масштабы увеличения, скопления частиц (кластеров), отдельные объекты и суб-части объектов сегментируются. Например, вирусы, имеющие доминирующий масштаб, могут быть достоверно

сегментированы. Поскольку сегментация выполняется в нескольких масштабах, регион с вирусом при определенном масштабе, в свою очередь, может быть также разделен на значимые подсегменты при более детальном масштабе увеличения, в то время как окружающие или фоновые участки вирусных частиц сегментируются при более грубом масштабе увеличения. Таким образом, нет необходимости априори определять, что на изображении представляет интерес, за исключением того, что каждое изображение имеет масштаб увеличения, чтобы все объекты в определенном масштабе были представлены наблюдателю изображения. Следует понимать, что изображение также включает в себя множество сегментов, которые не содержат вирусных частиц. Включение сегментов, не содержащих вирусных частиц, имеет значение для полноты онтологического исследования и может выявить неочевидные, но повторяющиеся структурные проявления.

На Фиг. 3 приведен схематический вид изображения 120 при более детальном масштабе увеличения, на котором показаны три вирусные частицы 122, 124, 126. Использование вирусных частиц является лишь наглядным примером применения принципов настоящего изобретения. Настоящее изобретение не ограничивается вирусными частицами, но может быть применено к любому объекту на изображении. Вирусная частица 122 в сегменте 128 содержит поверхностные белки 130, 132, 134, 136, 138 и 140 прикрепленные к ней. Эти белки расположены внутри сегментов 620, 622, 624, 626, 628 и 630, соответственно. Предпочтительно, чтобы сегментация поверхностных белков проводилась так же, как и для вирусных частиц, но в другом масштабе. Вирусная частица 124 в сегменте 142 имеет поверхностные белки 144, 146 и 148, прикрепленные к ней. Эти белки расположены внутри сегментов 638, 634 и 636, соответственно. Третья вирусная частица 126 в сегменте 150 не содержит прикрепленных белков. Изображение 120 также содержит свободный поверхностный белок 152 в сегменте 154, который не прикреплен ни к одной вирусной частице. Белок 152 может быть не заметен, если масштаб увеличения недостаточно крупный. Используя более детальные масштабы, способ настоящего изобретения может использовать особенности подструктур внутри или снаружи вирусной частицы, такие как поверхностные белки, которые могут быть прикреплены к вирусным частицам, чтобы найти дополнительные ассоциации. Способ также анализирует объекты, расположенные между вирусными частицами 122, 124, 126, путем анализа и определения яркости

каждого пикселя на изображении.

Следующим общим шагом математического алгоритма настоящего изобретения является преобразование, перестановка и перемещение идентифицированных объектов, расположенных внутри сегментов, в группы для более точного сравнения. Изменения размеров и формы объектов нормализуются путем исследования и использования собственных функций локального выравнивания, определяемых тем же способом, что и сегментация.

Как лучше всего показано на Фиг. 4, идентифицированные сегменты данных 128, 142 и 150 (содержащие частицы вируса 122, 124, 126 соответственно) преобразуются и выравниваются в пространстве, которое нормализует возможную вольность используемого способа сбора сигнала. В частности, для каждого идентифицированного сегмента рассчитываются и ортонормируются дополнительные локализованные собственные функции с полиномными коэффициентами вплоть до уровня N, что дает эффект векторных полей, которые согласуются со структурой изображения. Сегменты данных 128, 142, 150 преобразуются в сегменты 128', 142' и 150', которые можно рассматривать как имеющие новую систему координат, так что преобразованные объекты выровнены в сегментах 128', 142' и 150'. Все такие преобразованные сегменты сохраняются для дальнейшего анализа.

Теперь сегменты образуют новую функцию на сэмплированном изображении, а именно $I(s, x(y))$, которая сравнивается непосредственно под новой свободной переменной (x). Другим ключевым моментом настоящего изобретения является выбор и расчет локализованных собственных функций в каждом сегменте, которые выполняются при соблюдении условия ортогональности друг к другу.

Как схематично показано на Фиг. 4, каждый сегмент анализируется по отдельности для определения ориентации каждого сегмента и других нелинейных выравниваний, для которых он нормализован. Особенностью настоящего изобретения является то, что уравнение/алгоритм, используемый для определения сегментации, показанный на Фиг. 2A, 2B, также может быть использован для анализа сегментов после преобразования изображения. В частности, каждый сегмент, такой как сегмент 128, 142, 150, преобразуется до упрощенной формы, как показано стрелкой 155, и выравнивается в общей фиксированной ориентации, как показано на примере преобразованных сегментов 128', 142' и 150'. Каждый пиксель 600, 602, 604 и т.д. преобразованного сегмента

128' и всех остальных сегментов анализируется при первом масштабе для определения яркости пикселей. Затем значение яркости каждого пикселя сравнивается с соответствующими пикселями 606, 608, 610 и т.д. соответственно преобразованного сегмента 142' и всех остальных преобразованных сегментов.

- 5 Другими словами, значение яркости пикселя 600 сравнивается со значением яркости пикселя 606, пикселя 602 – с пикселем 608 и так далее. Это сравнение достаточно эффективно, так как преобразованные сегменты 128' и 142' имеют одинаковый размер, одинаковое количество пикселей и преобразованные объекты 122', 124' выровнены одинаково внутри каждого преобразованного
- 10 сегмента. Этот принцип применим ко всем сегментам, которые включают идентифицированные объекты, подструктуры и фон при любом масштабе увеличения. Следует отметить, что переход от базовой структуры прототипа к его первоначальному представлению может быть обращен.

- Преобразование и перестановка визуального впечатления от объектов
- 15 позволяют более точно сравнивать их. Объекты, имеющие одинаковую топологию или изо-структуру, после преобразования будут иметь одинаковую или похожую конфигурацию или форму, а направление объектов будет иметь ту же ориентацию. Например, преобразованные объекты могут иметь одинаковую длину и иметь верх и низ, выровненные друг с другом. Это показано на Фиг. 4, так как
- 20 преобразованные объекты 122', 124' и 126' имеют практически идентичные эллиптические формы, и они выровнены или направлены одинаково внутри сегментов 128', 142' и 150'. Если объекты представляют, например, один и тот же тип вирусных частиц, то преобразованные объекты 122', 124' и 126' также будут иметь ту же топологию. В этом случае светлый участок следует за темным
- 25 участком, за которым следует светлый участок. Перед преобразованием первый объект может иметь более яркую первую половину сегмента и более темную вторую половину сегмента, а второй - более яркую одну треть сегмента и более темные вторые две трети сегмента. Таким образом, топология может включать первый яркий сегмент, за которым следует второй темный сегмент. После
- 30 преобразования алгоритм будет, если используется одно измерение, изображать объекты таким образом, чтобы первая половина обоих объектов была яркой, а вторая половина - темной, хотя до преобразования второй объект имел 1/3 яркой и 2/3 темной. Если бы у объекта был первый темный сегмент, второй яркий сегмент и третий темный сегмент, то это составляло бы другую топологию.

- 35 На практике определение топологии может быть использовано для

идентификации объектов и подструктур, таких как вирусные частицы и поверхностные белки, прикрепленные к вирусным частицам. Следует отметить, что алгоритм не ограничивается одним измерением и что при преобразовании предпочтительно использовать несколько измерений. Это применимо даже в том случае, если частицы вируса перекрывают друг друга на изображении таким образом, что одна вирусная частица располагается чуть выше другой вирусной частицы.

Преобразование позволяет более точно сравнить преобразованные объекты и выявить дополнительные сходства или ассоциации, которые не могли быть видны или было очень трудно увидеть до преобразования. В общем, если объекты имеют одинаковую топологию, то после преобразования они будут выглядеть более похожими. Также можно анализировать "соседей" объектов. Например, один объект, такой как свободная вирусная частица, расположенный дальше от кластера, имеет соседние объекты, которые отличаются от вирусных частиц кластера. Хотя свободная вирусная частица очень похожа на вирусные частицы в кластере, она, скорее всего, имеет более высокую ассоциацию с другими соседними свободными частицами, чем с вирусными частицами в кластере. Также можно анализировать подструктуры вирусных частиц, т.е. проводить анализ с более крупным масштабом увеличения.

Как указывалось ранее, вирусные частицы могут иметь поверхностные белки, которые могут быть проанализированы для определения уровня ассоциации между вирусными частицами и поверхностными белками. Не имеет значения, где на вирусных частицах прикреплены поверхностные белки, поскольку основное внимание в исследовании уделяется типу и количеству белков, которые прикреплены к вирусной частице. Можно также проанализировать обстановку или фон, окружающий вирусную частицу, чтобы усилить или ослабить ассоциацию между вирусными частицами в каждом сегменте.

Анализ преобразованного изображения также является автоматизированным процессом, в котором алгоритм в компьютере проходит через каждый пиксель в преобразованном сегменте для определения уровня яркости каждого пикселя. Человеческие решения о том, как изменить форму вирусной частицы или какие частицы следует сравнивать друг с другом, либо отсутствуют, либо очень ограничены. Как показано на Фиг. 4, преобразование изображения вирусной частицы 122 в сегменте 128 может выпрямить и

растянуть/укоротить его до заданной длины, как показано на преобразованной вирусной частице 122' в сегменте 128'. Аналогично, вирусная частица 124 выпрямляется до вирусной частицы 124' в сегменте 142', который предпочтительно имеет примерно такую же длину и выровнена также как и вирусная частица 122', так, чтобы и верх и низ были выровнены. То же преобразование вирусной частицы 126 до вирусной частицы 126' происходит в сегменте 150', которая выравнивается с вирусными частицами 122' и 124'. Частицы 122', 124' и 126' теперь можно сравнить более точно. Частицы 122, 124 и 126 имеют одинаковую топологию, поэтому после преобразования частицы 122', 124' и 126' имеют очень похожую форму или конфигурацию. Анализируются также "соседи" вирусных частиц, т.е. поверхностные белки. Аналогичным образом, изображения поверхностных белков преобразуются в поверхностные белки 130', 132', 134', 136', 148', 144', 146', 148 и 152' соответственно.

Следует отметить, что позиции поверхностных белков, которые прикреплены к вирусным частицам или находятся в непосредственной близости от них, отличаются. Если бы поверхностные белки были расположены в одном и том же месте на каждой частице вируса, то поверхностные белки можно было бы включить в сравнение при первом масштабе для каждого пикселя, например, между преобразованными сегментами 128' и 142', как часть вирусных частиц. Однако, поскольку поверхностные белки расположены в разных местах, способ настоящего изобретения также проводит анализ на втором более детальном масштабе, чтобы сначала определить существование белков на предварительном этапе сегментации (не обращая внимания на точное местоположение поверхностных белков). На более детальном масштабе, например, на втором, компьютер затем проходит через него и сравнивает каждый пиксель в преобразованных сегментах для поверхностных белков и всех других преобразованных сегментов на разных уровнях, чтобы, например, определить топологии поверхностных белков, которые прикреплены или находятся вблизи от вирусных частиц. Преобразованные поверхностные белки представлены в виде 130', 132', 134', 136', 138', 140' и 152'. Преобразованные поверхностные белки еще не отсортированы, а только сравниваются друг с другом в таком масштабе, который подходит для анализа поверхностных белков. Аналогично анализу каждого пикселя в преобразованных сегментах 128', 142' и 150' анализируется каждый пиксель для преобразованных сегментов 620', 622', 624', 626', 628', 630', 154', 634', 636' и 638', но, например, с большим увеличением или более

детальным масштабом. Таким образом, значение яркости каждого пиксела, например, преобразованного сегмента 620' сравнивается со значениями яркости соответствующих пикселей преобразованного сегмента 622' и всех остальных преобразованных сегментов. Другими словами, первый пиксел в левом верхнем углу преобразованного сегмента 620' сравнивается с соответствующим первым пикселом в левом верхнем углу преобразованного сегмента 622' и так далее, пока уровень яркости всех пикселей не будет сравниваться со всеми соответствующими пикселями в других преобразованных сегментах. Эти значения сохраняются компьютером для дальнейшего анализа и определения топологии.

Важной особенностью настоящего изобретения является то, что инструкция по перестановке встроена в сами объекты (т.е. ассоциации между объектами) таким образом, что после преобразования и сравнения объекты с наиболее сильной ассоциацией перемещаются в одну и ту же группу. Она не должна определяться человеком или оптимизировать расположение между каждой парой изображений.

Ниже описан принцип использования автоматизированных ассоциаций между объектами в различных масштабах и пространственной близости. Первый масштаб увеличения может быть пригоден для анализа кластера вирусных частиц, второй масштаб увеличения может быть наиболее подходящим для анализа отдельных вирусных частиц, а третий масштаб увеличения может быть пригоден для анализа подструктур внутри и вне вирусных частиц. Повторяющиеся объекты при разных масштабах идентифицируются как группы, которые дают пользователю представление о том, какие масштабы увеличения представляют интерес для определенного анализа.

Разработанный способ автоматизированной ассоциации сегментов данных при различных масштабах и пространственной близости может быть математически описан, как показано ниже. Извлеченные сегменты помещаются в конечный граф с ядром ассоциативности.

$$g_2(c1, c2) = \phi(k_1 d_s(c_1, c_2)) + k_2 \sum_{a \in n_{c1}}^{\min} ||p_a - p_b||^2$$

Где расстояние d_s между двумя преобразованными сегментами является стандартной нормой L^2 для разницы функций, а соседние сегменты обозначаются (n_{c1}) и (n_{c2}) . Собственные векторы для этого ядра ортонормированы до полиномиального уровня 3 и предназначены для кластеризации. Добавленная

точка собственного вектора сегмента (S) называется (p).

Как показано на Фиг. 5, преобразованные объекты, такие как вирусные частицы, другие под-структуры и фоновые сегменты с высокой ассоциацией друг к другу группируются вместе. Как описано выше, значение яркости каждого пикселя в преобразованных сегментах сравнивается с соответствующим пикселем во всех других преобразованных сегментах. В общем, преобразованные объекты внутри преобразованных сегментов, похожие на другие преобразованные объекты, перемещаются ближе друг к другу или сгруппированные вместе.

Другими словами, при первом масштабе значения пикселей для каждого преобразованного сегмента, такого как сегменты 122', 150' и 142' на Фиг. 5 сравниваются друг с другом и со всеми другими сегментами, так что значения пикселей в одном преобразованном сегменте сравниваются со значениями пикселей в другом преобразованном сегменте для определения похожей или той же топологии. Сегменты с наибольшей ассоциацией или схожестью группируются вместе. Компьютер присваивает каждому преобразованному сегменту пространственную координату таким образом, чтобы аналогичные преобразованные сегменты получали пространственные координаты, которые являются схожими, и поэтому преобразованные сегменты перемещаются ближе друг к другу. При втором масштабе, который детальнее, чем первый, значения пикселей для каждого сегмента преобразованных сегментов сравниваются друг с другом, так что преобразованные изображения белков, имеющие одинаковую топологию, группируются вместе, как показано на Фиг. 5. Другими словами, преобразованные объекты, имеющие наибольшую ассоциацию, перемещаются ближе всего друг к другу в группе. Предпочтительно, чтобы группировка осуществлялась в различных масштабах. Анализируемые сегменты могут быть не из одного изображения. Преобразованные частицы 122', 124' и 126' группируются вместе, поскольку они имеют более высокую ассоциацию друг с другом, т.е. одинаковую или схожую топологию, чем другие объекты на изображении. Аналогично, поверхностные белки 130', 136', 146', 148' и 152' (проиллюстрированные как "Т-белки") группируются вместе, поскольку они имеют одинаковую или сходную топологию, поверхностные белки 134', 140' и 144' (проиллюстрированные Y-белки) группируются вместе и поверхностные белки 132' и 138' (проиллюстрированные как I-белки) группируются вместе. Все преобразованные сегменты сравниваются друг с другом. Это означает, например, что пиксели в преобразованном сегменте 128' сравниваются с пикселями в

преобразованном сегменте 620' (который содержит преобразованный поверхностный белок 130') и поскольку топологии сильно отличаются, пространственные координаты, назначенные этим двум сегментам компьютером, сильно отличаются. Таким образом, сегмент 620' находится достаточно далеко от сегмента 128', как показано на Фиг. 5.

Как поясняется ниже, каждый сегмент подобных объектов в группировке может быть затем проанализирован на более высоком уровне ассоциации для выявления дополнительных и более детальных/специфических ассоциаций, как показано на Фиг. 6, что создает онтологические древовидные структуры. На Фиг. 5 сегменты организованы и отсортированы по группам в зависимости от того, включают ли они объекты или структуры со сходной или одинаковой топологией. Например, вирусные частицы ассоциируются или группируются вместе, потому что каждый сегмент содержит вирусные частицы. Кроме того, сходные поверхностные белки группируются вместе на основе прямого сравнения. Однако соседние сегменты могут содержать дополнительную информацию, которая превращает их в новые повторяющиеся комбинации сегментов. На Фиг. 6 сегменты сравниваются друг с другом для определения наличия общих пикселей в представлении предварительного преобразования (см. Фиг. 3), чтобы найти дополнительные ассоциации. Сосед" может быть определен как сегмент, содержащий подструктуру, которая одновременно имеет ту же площадь, что и сегмент, содержащий вирусную частицу, т.е. сегмент подструктуры находится внутри сегмента вирусной частицы, так что сегмент вирусной частицы охватывает сегмент поверхностного белка, как показано на Фиг. 3. Одна и та же область означает, что два сегмента на Фиг. 3 имеют по крайней мере один пиксель одновременно в двухмерном представлении, например, в плоскости xy. Таким образом, способ определяет, являются ли два сегмента соседними или нет, определяя, имеют ли оба сегмента по крайней мере один общий пиксель. Это не определяет точное положение поверхностного белка или вирусной частицы в сегменте вирусной частицы. Способ просто фокусируется на том, есть ли общие пиксельные области для этих двух сегментов. Как указывалось выше, информация о пикселях, используемая для сегментации, как показано на Фиг. 3, сохраняется в базе данных для извлечения с целью анализа, являются ли два сегмента соседними или нет. Следует отметить, что вирусные частицы могут быть проанализированы при первом масштабе, а подструктуры (т.е. поверхностные белки) - при втором. Если масштаб увеличения рассматривается как третье

измерение, и две частицы вируса находятся в одном масштабе увеличения, то две частицы вируса могут рассматриваться как соседние, хотя они могут и не быть соседями в плоскости ху. Таким образом, два сегмента могут быть соседями или соседями в пространстве, поскольку сегменты анализируются при одном и том же масштабе увеличения.

На рис. 6 проиллюстрирована мысль о том, что наличие поверхностных белков мелкого масштаба и их индивидуальная ассоциация дают дополнительную информацию об ассоциации сложных вирусных частиц 122', 124' и 126'. Теперь компьютер анализирует соседей вирусных частиц, т.е. подструктуры, которые были идентифицированы на этапе сегментации внутри каждого сегмента вирусной частицы таким образом, что сегменты подструктур имеют по крайней мере один общий пиксель с сегментом вирусной частицы. Как показано на Фиг. 3, несколько сегментов, содержащих поверхностные белки, имеют ту же площадь, что и сегменты, содержащие вирусные частицы, т.е. сегменты, содержащие поверхностные белки, расположены внутри сегмента вирусных частиц. Вместо того, чтобы просто сравнивать каждый пиксель трансформированных сегментов друг с другом, как показано на Рис. 5, сегменты (содержащие подструктуры, т.е. поверхностные белки), которые занимают ту же площадь, что и сегмент, содержащий вирусную частицу, сравниваются с сегментами (содержащими подструктуры), которые занимают ту же площадь, что и сегмент другой вирусной частицы. Таким образом, анализируются сегменты поверхностных белков, которые занимают участки/пиксели с другими сегментами, например, сегменты, содержащие частицу вируса. Точное положение сегмента подструктуры внутри более крупного сегмента не анализируется и не определяется. Например, существует более сильная ассоциация между преобразованными вирусными частицами 122' и 124', каждая из которых имеет поверхностные белки в качестве соседей внутри преобразованных сегментов 128', 142', соответственно, по сравнению с ассоциацией с преобразованной вирусной частицей 126', которая не имеет поверхностных белков, расположенных внутри ее преобразованного сегмента 150', т.е. на Фиг.3 внутри сегмента 150, окружающего вирусную частицу 126, не было сегмента, содержащего поверхностный белок. Таким образом, преобразованным сегментам 128' и 142' компьютер присваивает новые более точные пространственные координаты так, чтобы они перемещались ближе друг к другу, чем пространственные координаты, присваиваемые преобразованному сегменту 150' (содержащему вирусную частицу 126') к преобразованным

сегментам 128', 142'.

Аналогичным образом, существование вирусных частиц в качестве соседей поверхностных белков дает дополнительную информацию об ассоциации между поверхностными белками таким образом, что это обеспечивает более сильную ассоциацию поверхностных белков, которые прикреплены к вирусной частице, чем "свободные" поверхностные белки, которые не прикреплены ни к одной вирусной частице. В частности, преобразованный сегмент 154' (содержащий преобразованный поверхностный белок 152') не занимает общей площади с сегментом, содержащим вирусную частицу. Поверхностный белок 152' - это "свободный" поверхностный белок, который не прикреплен к вирусной частице, в то время как сегменты, содержащие поверхностные белки 130', 136', 144' и 146' все занимают область с сегментами, каждый из которых содержит вирусную частицу, и поэтому между сегментами этих поверхностных белков существует более высокая связь, чем между сегментами 154', содержащими только свободный белок 152'. Другими словами, на разных шкалах восприятия сегменты, таким образом, могут быть отсортированы и разделены на множества в "перцептивном" пространстве. Каждый набор сегментов можно далее разделить на подмножества, если некоторые из сегментов имеют более сильную ассоциацию друг с другом, чем другие. Важным аспектом настоящего изобретения является то, что ассоциация является инклюзивной без выбора различающих свойств, а только с инклюзивными свойствами. Единственным разделяющим свойством является общая нормализация, заставляющая точки данных распространяться по единичному шару.

Ниже описан способ автоматизированной проверки гипотез ассоциативных кластеров. Способ начинается с мелкого масштаба, обрабатываемого уравнениями (с1) и (с2) (см. уравнения выше), в которых ядро имеет только одну фиксированную точку. Постепенное приближение к более крупному масштабу увеличения, где все сегменты уникальны и отдельная группа вокруг сегмента (S) оценивается с помощью шариков вокруг сегментов. Кластер определяется если:

$$\{x \in \Sigma, x: r < \|p_s - p_x\| < 2r\} = \emptyset \text{ для некоторых } (r).$$

Дополнительно, то же будет сохраняться для каждой внутренней точки в $I = \{x \in \Sigma, x: \|p_s - p_x\| < r\}$ потенциального кластера с радиусом (r), то есть $\forall i \in I: \{x \in \Sigma, x: r < \|p_i - p_x\| < 2r\} = \emptyset$, определяющее локальную единственность предлагаемого кластера. Кластер не должен быть пустым и должен иметь непустое множество внешних точек (расположенных на расстоянии $> 2r$).

Таким образом, это автоматизированный способ, который не требует никаких предположений о данных для сортировки данных по разделяемым классам. Таким образом, настоящее изобретение позволяет производить автоматизированные расчеты объединения объектов и онтологические исследования многих объектов друг с другом оперативно и энергоэффективно, без необходимости делать сложные предположения, как это происходит при использовании заранее определенных фильтров и без необходимости оптимизации набора параметров при сравнении двух объектов.

На Фиг. 7 схематично показано, как вирусные частицы и поверхностные белки кластеризуются в группы (как показано пунктирными кругами), исходя из того, является ли группировка значительной, как это определено по результатам измерения ассоциации. Разделение множества на подмножества может происходить, когда подмножество структур расположено на значительном расстоянии от всех других точек подмножества. Например, вирусная частица 126', не содержащая поверхностных белков, расположена дальше в "перцептивном" пространстве от двух вирусных частиц 122' и 124', которые содержат поверхностные белки, по сравнению с расстоянием между двумя вирусными частицами 122' и 124' (которые содержат поверхностные белки). Однако вирусные частицы 122', 124' и 126' объединены в группу 200 вирусных частиц, в то время как вирусные частицы 122' и 124' объединены в группу 202 вирусных частиц, к которым прикреплены поверхностные белки. Аналогично, поверхностные белки 130', 136', 144', 146' и 152' также объединены в группу 204 белка типа "Т". Только поверхностные белки 130', 136', 144' и 146' сгруппированы в группу 206 белков "Т", которые прикреплены к вирусной частице. Поверхностные белки 134', 140' и 148' объединены в группу 208' для белков "Y", а поверхностные белки 132' и 138' - в группу 210 для белков "I".

Кластеризация, описанная на Фиг. 7, аналогична этапу сегментации, показанному на Фиг. 3, где объекты и подструктуры, имеющие наибольшую ассоциацию (например, на основе топологии и критериев соседства), группируются в общий кластер или значительные/четкие группы, так что чем ближе сегменты расположены в пространстве сортировки (на основе координат и собственных векторов), тем выше связь между ними. Различающимися группами могут быть группы, которые находятся далеко друг от друга в системе координат, и между группами нет или очень мало объектов/структур.

Ядро ассоциации g_2 может быть дополнено дополнительными условиями

ассоциации между сегментами и группами сегментов. Как показано на Фиг. 9-11, соотношение между сегментом с двумя выбранными соседними сегментами можно сравнить с соотношением между другим сегментом с двумя выбранными соседними сегментами. Такое сравнение отношений выбирается для измерения

5 расстояний между тремя объектами в группах (т.е. сегментом и двумя соседями) в виде треугольника. Два сегмента, которые имеют схожие треугольные отношения неравенства с соседями, имеют высокую ассоциацию. Реальным примером такого когнитивного сравнения является соотношение между электроном и ядром атома в контексте атома по сравнению с соотношением между Землей и Солнцем в

10 контексте Солнечной системы. В обеих системах есть, например, центральный объект, вокруг которого небольшие объекты движутся по овальному или круговому пути, поэтому отношения или взаимодействия между компонентами в обеих системах схожи.

Два идентифицированных кластера могут пройти ту же локальную схему,

15 что и описанная выше. Распределение позиций сегмента в каждом локальном кластере описывает локальную морфологическую геометрию группы сегментов. Эту морфологическую геометрию можно сравнить между двумя кластерами сегментов для того, чтобы группировать кластеры в зависимости от их группового поведения. Морфологическая геометрия - это, например, изменение объекта,

20 рассматриваемого с разных позиций, или небольшое, но непрерывное изменение топологии объекта. Один из способов сравнения двух распределений, предпочтительно, основан на сумме наименьших расстояний между точками в каждом кластере.

В частности, на Фиг. 9 показан такой объект, как вирусная частица 222 в

25 сегменте 228, который во многом похож на вирусную частицу 122, показанную на Фиг. 3-4. Хотя частица 222 похожа на частицу 122, все детали или подструктуры, прикрепленные к частицам, отличаются или очень сильно отличаются от деталей частицы 122, но соотношение между деталями остается тем же или очень похожим. Для простоты здесь используются термины "детали", "подструктура" и

30 "поверхностные белки", которые являются взаимозаменяемыми.

Поверхностные белки 230, 232, 234, 236, 238 и 240 отличаются от поверхностных белков 130, 132, 134, 136, 138 и 140 частицы 122, но взаимоотношения между поверхностными белками схожи или одинаковы, например, их относительное положение по отношению к другим поверхностным

35 белкам. Следует понимать, что детали 230-240 не обязательно должны быть

поверхностными белками и могут быть любым объектом, отличным от деталей 130-140. Следует также понимать, что расположение деталей 230-240 на частице 222 не обязательно должно совпадать с расположением поверхностных белков 130-140 на частице 122, поскольку именно взаимосвязь между деталями или

5 поверхностными белками представляет наибольший интерес. Если в приведенном выше примере детали 230-240 и детали 130-140 окажутся также расположены в одном и том же положении каждой частицы, соответственно, это еще больше усиливает сходство того, как детали соотносятся друг с другом. Следует понимать, что точное пространственное положение деталей 130-140,

10 230-240 на каждой соответствующей частице 122, 222 не является основной целью. Вместо этого, основное внимание может быть сосредоточено на том, как один поверхностный белок расположен относительно другого поверхностного белка, не зная точно, где на частице прикреплены поверхностные белки. Например, Y-деталь 234' имеет I-деталь 232' слева и T-деталь 236' справа, что

15 аналогично Y-детали 134', которая имеет I-деталь 132' слева и T-деталь 136' справа от Y-детали 134' также, как показано на Фиг. 10-11.

Также возможно, что детали 230-240 также являются поверхностными белками, которые идентичны деталям 130-140, но детали 230-240 имеют дополнительные вторичные поверхностные белки, прикрепленные к

20 поверхностным белкам, которые аналогичны поверхностным белкам 130-140. Это отличает поверхностные белки 230-240 от поверхностных белков 130-140, но они связаны друг с другом и с основной частицей/вирусом одинаково, поэтому группировки вирусных частиц и поверхностных белков, прикрепленных к ним, связаны друг с другом аналогичным или одинаковым образом.

25 Новейшая концепция сравнения неравенства треугольников в настоящем изобретении измеряет это сходство и установлено, что частица 222 и частица 122 более тесно связаны друг с другом, чем частица 122 и частица 124 (что было показано на рис. 3-4), несмотря на то, что частица 122 и частица 124 имеют некоторые общие детали, т.е. одни и те же поверхностные белки. Другими

30 словами, есть "прямое сходство" между частицами 122 и 124, поскольку они содержат одни и те же поверхностные T-белки и поверхностные Y-белки. Однако выявление прямых сходств между объектами может быть весьма затруднительным, поскольку необходимо провести так много испытаний характеристик объектов. В сравнении неравенства треугольников акцент делается

35 не на прямых сходствах между деталями (т.е. поверхностными белками) частиц, а

скорее на том, как детали связаны друг с другом или как детали взаимодействуют друг с другом, чтобы найти или определить связи, которые имеют общие черты, даже если детали не похожи друг с другом, когда напрямую сравниваются. Например, связь между поверхностными белками, которые прикреплены к вирусной частице в определенном контексте, может быть перенесена на другие поверхностные белки, которые также прикреплены к вирусной частице в другом контексте, чтобы определить общие связи в обоих контекстах.

В частности, белки 130, 136 частицы 122 идентичны белкам 146, 148 частицы 124, а поверхностные белки 134, 140 частицы 122 идентичны поверхностным белкам 144 частицы 124. Однако соотношение между поверхностными белками отличается для частицы 124 по сравнению с частицей 122. Кроме того, частица 124 полностью лишена поверхностных белков, которые аналогичны поверхностным белкам 132, 136 частицы 122. Например, частица 124 не содержит деталей типа "I". Это показано на Фиг. 10-11 (аналогично иллюстрированному переходу на Фиг. 4-5).

Фиг. 10 по существу похожа на Фиг. 7, но преобразованная частица 222' и ее детали 230'-240' были добавлены для проведения сравнения неравенства треугольников с частицами 122' и 124'. Близость или соседство на Фиг. 10 представляет собой схожесть между объектами. Как лучше всего показано на Фиг. 11, линия 400а показывает, как преобразованная деталь 236' соотносится с преобразованной деталью 234' преобразованной частицы 222'. Например, длина, положение и направление линии 400а могут указывать на эту взаимосвязь, поэтому эта геометрия используется для иллюстрации взаимосвязи деталей друг с другом. Могут быть использованы и другие характеристики линии. Аналогичным образом, линия 402а показывает, как преобразованная деталь 234' соотносится с преобразованной деталью 232'. Линия 400b показывает, как преобразованная деталь (например, поверхностный белок) 136' связана с преобразованной деталью 134' преобразованной частицы 122' и линия 402b, показывающая как преобразованная деталь 134' связана с преобразованной деталью 132' преобразованной частицы 122'. Можно было бы проиллюстрировать взаимосвязь между деталью 236' и деталью 238', продолжив другую линию непосредственно от детали 236' и 238' так, чтобы линии 400а, 400b и новая линия вместе образовывали треугольник. В этом случае, поскольку линии 400а и 402а вместе образуют прямую линию, это означает, что расстояние между деталью 236' и деталью 234' плюс расстояние между деталью 234' и деталью 232' совпадает с

расстоянием между деталью 236' и деталью 232' и что эти три линии образуют треугольник, не имеющий высоты, а линия между деталью 236' и деталью 232' аналогична линии 400а плюс линия 402а. Расстояние между деталью 236' и деталью 232' могло бы быть короче, и линии 400а и 402а не образовывали бы

5 прямую линию, но угол между ними образуется таким образом, что три линии образуют видимый треугольник, имеющий высоту. Линия 404 показывает, как преобразованная деталь 146' преобразованной частицы 124' (которая отличается от детали 122') соотносится с преобразованной деталью 144' той же частицы 124'. Как видно, линия 400а по существу аналогична линии 400b и имеет примерно

10 одинаковую длину и направление, что означает, что связь между деталью 236' и деталью 234' (представленная линией 400а) по существу аналогична связи детали 136' с деталью 134' (представленная линией 400b). Линия 404, с другой стороны, имеет длину и направление, более отличное, чем линия 400а имеет к линии 400b, так что даже если бы частица 124' имела I-детали, образовавшийся

15 треугольник отличался бы от треугольника, образовавшегося между деталями 236'-234'-232' (треугольник без высоты) и 136'-134'-132' (также треугольник без высоты, совпадающий с треугольником 236'-234'-232', расположенным выше). Длина и направление линий зависят от всех деталей в контексте, поэтому если, например, детали 232' и 238' отсутствовали, то частица 222' могла быть более

20 похожа на частицу 124' вместо того, чтобы быть более похожей на частицу 122'.

Поскольку частица 122' и детали или белки поверхности 130'-140', присоединенные к ней, больше похожи на частицу 222', а детали или белки поверхности 230'-240', присоединенные к ней, чем к частице 122' (включая детали 130'-140'), относятся к частице 124' (включая детали 134', 140' и 144'), то частица

25 122' движется ближе, как показано стрелкой 250, к частице 222', образуя групповое скопление 252', так что скопление частицы 222' ближе к скоплению частицы 122', чем скопление частицы 124', к скоплению частицы 222'. Измерение неравенства треугольников между 222', 122' и 124' скоплениями частиц таково, что расстояние D1 между частицами 222' и 122' частицами является наименьшим,

30 тогда как расстояния D2, D3 от частицы 122' до частицы 124' и от частицы 222' до частицы 124', соответственно, больше, как показано внизу на Фиг. 11.

Хотя настоящее изобретение было описано в соответствии с предпочтительными композициями и вариантами осуществления, следует понимать, что некоторые замены и изменения в нем могут быть сделаны без

35 отклонения от духа и объема следующих пунктов формулы.

ФОРМУЛА

1. Способ автоматизированной онтологической группировки объектов, содержащий:

получение цифровой камерой (300) снимка изображения (302) в электронном
5 микроскопе (304);

сохранение изображения в виде цифрового изображения (100) в компьютере (306);

при первом масштабе увеличения компьютер автоматически идентифицирует
объекты (102, 122, 124, 126) на цифровом изображении (100, 120) путем
10 сравнения яркости пикселей (600, 602, 604, 606, 608, 610) на цифровом изображении;

определение компьютером степени ассоциации между каждым пикселем на основе яркости каждого пикселя;

формирование компьютером первого сегмента (128) вокруг первого
15 идентифицированного объекта (122) и второго сегмента (142) вокруг второго идентифицированного объекта (124) на цифровом изображении (120), разделенными, где пиксели имеют наименьшую связь друг с другом;

математическое преобразование компьютером первого идентифицированного объекта в первый преобразованный объект (122') в первом преобразованном
20 сегменте (128') и второго идентифицированного объекта во второй преобразованный объект (124') во втором преобразованном сегменте (142'), при этом преобразованные первый и второй объекты (122', 124') имеют формы с фиксированной ориентацией;

определение первой топологии преобразованного первого объекта (122')
25 компьютером на основе яркости пикселей; и

сравнение компьютером яркости пикселей (600, 602, 604) в первом преобразованном сегменте (128') с яркостью пикселей (606, 608, 610) во втором преобразованном сегменте (142').

30 2. Способ по пункту 1, в котором способ дополнительно содержит шаги при втором масштабе увеличения, причем второй масштаб увеличения более крупный, чем первый масштаб увеличения, на которых автоматически идентифицируют компьютером подструктуры цифрового изображения путем сравнения яркости пикселей на цифровом изображении;

35 определяют компьютером степень ассоциации между каждым пикселем в

зависимости от яркости каждого пикселя;

формируют компьютером первый сегмент подструктуры (620, 622, 624, 626, 628, 630) вокруг первой идентифицированной подструктуры (130, 132, 134, 136, 138, 140) и второй сегмент подструктуры (634, 636, 638) вокруг второй идентифицированной подструктуры (144, 146, 148) в цифровом изображении, разделенными, где пиксели имеют наименьшую связь друг с другом;

математическое преобразование компьютером первой идентифицированной подструктуры в первую преобразованную подструктуру (130', 132', 134', 136', 138', 140') в первом преобразованном сегменте подструктуры (620', 622', 624', 626', 628', 630') и вторую идентифицированную подструктуру во вторую преобразованную подструктуру (144', 146', 148') во втором преобразованном сегменте подструктуры (634', 636', 638'), в котором преобразованные первая и вторая подструктуры имеют формы с фиксированной ориентацией;

определение второй топологии преобразованной первой подструктуры компьютером на основе яркости пикселей; и

сравнение компьютером яркости пикселей в первом преобразованном сегменте подструктуры с яркостью пикселей во втором преобразованном сегменте подструктуры.

3. Способ по пункту 2, в котором способ дополнительно содержит этапы определения компьютером первой топологии второго преобразованного объекта (124') и определения второй топологии второй преобразованной подструктуры (144', 146', 148'), на основе топологий компьютер присваивает координаты первому преобразованному объекту (122') и второму преобразованному объекту (124') таким образом, чтобы первый преобразованный объект и второй преобразованный объект находились в непосредственной близости в перцептивном пространстве и присваивает координаты первой преобразованной подструктуре, и второй преобразованной подструктуре таким образом, чтобы первая преобразованная подструктура и вторая преобразованная подструктура находились в непосредственной близости в перцептивном пространстве, но на удалении от первого и второго преобразованных объектов.

4. Способ по пункту 3, в котором способ дополнительно содержит этапы идентификации компьютером третьего объекта (126) и формирования третьего сегмента (150) вокруг третьего объекта (126), определяя, что первый сегмент

(128) и первый сегмент подструктуры (620) имеют общий пиксел и что второй сегмент (142) и вторая подструктура (636) имеют общий пиксел, причем компьютер определяет, что третий сегмент (150) не имеет общего пиксела с первой и второй подструктурами (620, 636), при этом компьютер назначает координаты первому и второму сегментам (128, 142) таким образом, что первый и второй сегменты ближе друг к другу в перцептивном пространстве, чем третий сегмент (150) к первому и второму сегментам (128, 142).

5. Способ по пункту 2, в котором способ дополнительно содержит этапы идентификации компьютером объектов при первом масштабе увеличения одновременно с идентификацией подструктур при втором масштабе увеличения.

6. Способ по пункту 3, в котором способ дополнительно содержит этапы, на которых идентифицируют третий объект (150) на втором цифровом изображении.

7. Способ по пункту 1, в котором способ дополнительно содержит этап, на котором идентифицируют четвертый объект (222), к которому прикреплены подструктуры (230, 232, 234, 236, 238, 240), преобразуют объект (222) и подструктуры (230-240) в объект (222') и подструктуры (230'-240'), при сравнении неравенства треугольников анализируют как подструктуры (230'-240') соотносятся друг с другом по сравнению с тем, как подструктуры (130'-140') соотносятся друг с другом и как подструктуры (144', 146' и 148') соотносятся друг с другом, для определения сходств.

ИЗМЕНЕННАЯ ФОРМУЛА ИЗОБРЕТЕНИЯ

получена Международным бюро 14 ноября 2018 года (14.11.2018)

1. Способ автоматизированной онтологической группировки объектов,
5 содержащий:
получение цифровой камерой (300) снимка изображения (302) в электронном микроскопе (304);
сохранение изображения в виде цифрового изображения (100) в компьютере (306);
- 10 при этом цифровое изображение (100) имеет последовательность пикселей со значениями яркости, содержащую первый пиксел, имеющий первое значение яркости, второй пиксел, смежный с первым пикселом, имеющий второе значение яркости, третий пиксел, смежный со вторым пикселом, имеющий третье значение яркости, и четвертый пиксел, смежный с третьим пикселом, имеющий четвертое
15 значение яркости, при этом цифровое изображение (100) также имеет пятый пиксель, имеющий пятое значение яркости, шестой пиксель, смежный с пятым пикселем, имеющий шестую яркость, седьмой пиксель, смежный с шестым пикселем, имеющий седьмую яркость, и восьмой пиксель, смежный с седьмым пикселем, имеющий восьмую яркость,
- 20 автоматическое определение компьютером, одновременно используя все возможные масштабы увеличения электронного микроскопа (304), степень ассоциации между пикселями, сравнивая первое значение яркости со вторым значением яркости, второе значение яркости с третьим значением яркости, третье значение яркости с четвертым значением яркости, компьютер также сравнивает
25 пятое значение яркости с шестым значением яркости, шестое значение яркости с седьмым значением яркости, седьмое значение яркости с восьмым значением яркости;
- формирование компьютером первого сегмента (128) вокруг первого идентифицированного объекта (122) между вторым и третьим пикселом, когда
30 второе значение яркости больше сопоставимо с первым значением яркости, чем второе значение яркости с третьим значением яркости, и когда третье значение яркости больше сопоставимо с четвертым значением яркости, чем третье значение яркости с вторым значением яркости;
- одновременно с формированием первого сегмента (128), формирование
35 компьютером второго сегмента (142) вокруг второго идентифицированного

объекта (124) на цифровом изображении (120) между шестым и пятым пикселом, когда шестое значение яркости больше сопоставимо с пятым значением яркости, чем шестое значение яркости с седьмым значением яркости, и когда седьмое значение яркости больше сопоставимо с восьмым значением яркости, чем седьмое значение яркости с шестым значением яркости;

математическое преобразование компьютером первого идентифицированного объекта в первый преобразованный объект (122') в первом преобразованном сегменте (128') и второго идентифицированного объекта во второй преобразованный объект (124') во втором преобразованном сегменте (142'), при этом преобразованные первый и второй объекты (122', 124') имеют формы с фиксированной ориентацией;

определение компьютером первой топологии преобразованного первого объекта (122') и второй топологии второго объекта (124') сравнивая яркости пикселей (600, 602, 604) в первом преобразованном сегменте (128') с яркостью соответствующих пикселей (606, 608, 610) во втором преобразованном сегменте (142'); и

присвоение компьютером пространственных координат первому преобразованному объекту (122') и второму преобразованному объекту (124') на основе выявленных топологий таким образом, что первый преобразованный объект (122') перемещается непосредственно ко второму преобразованному объекту (124'), когда первая топология аналогична второй топологии, и на удалении от второго преобразованного объекта (124'), когда первая топология отличается от второй топологии.

2. Способ по пункту 1, в котором способ дополнительно содержит шаги при втором масштабе увеличения, причем второй масштаб увеличения более крупный, чем первый масштаб увеличения, на которых автоматически идентифицируют компьютером подструктуры цифрового изображения путем сравнения яркости пикселей на цифровом изображении;

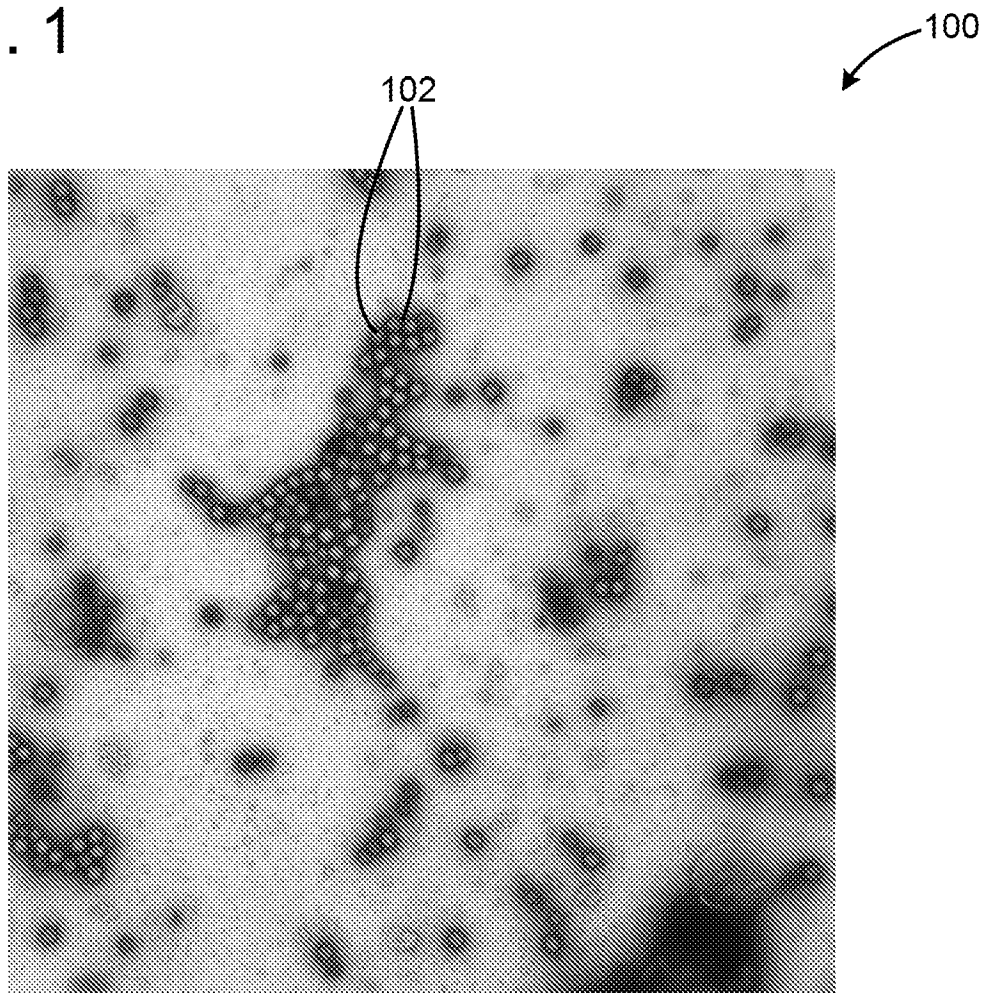
определяют компьютером степень ассоциации между каждым пикселем в зависимости от яркости каждого пикселя;

формируют компьютером первый сегмент подструктуры (620, 622, 624, 626, 628, 630) вокруг первой идентифицированной подструктуры (130, 132, 134, 136, 138, 140) и второй сегмент подструктуры (634, 636, 638) вокруг второй идентифицированной подструктуры (144, 146, 148) в цифровом изображении, разделенными, где пиксели имеют наименьшую связь друг с другом;

- математическое преобразование компьютером первой идентифицированной подструктуры в первую преобразованную подструктуру (130', 132', 134', 136', 138', 140') в первом преобразованном сегменте подструктуры (620', 622', 624', 626', 628', 630') и вторую идентифицированную подструктуру во вторую преобразованную подструктуру (144', 146', 148') во втором преобразованном сегменте подструктуры (634', 636', 638'), в котором преобразованные первая и вторая подструктуры имеют формы с фиксированной ориентацией;
- определение второй топологии преобразованной первой подструктуры компьютером на основе яркости пикселей; и
- сравнение компьютером яркости пикселей в первом преобразованном сегменте подструктуры с яркостью пикселей во втором преобразованном сегменте подструктуры.
3. Способ по пункту 2, в котором способ дополнительно содержит этапы
- определения компьютером топологии второй преобразованной подструктуры (144', 146', 148'), на основе топологий компьютер присваивает координаты первой преобразованной подструктуре и второй преобразованной подструктуре таким образом, чтобы первая преобразованная подструктура и вторая преобразованная подструктура находились в непосредственной близости в перцептивном пространстве, но на удалении от первого и второго преобразованных объектов.
4. Способ по пункту 3, в котором способ дополнительно содержит этапы идентификации компьютером третьего объекта (126) и формирования третьего сегмента (150) вокруг третьего объекта (126), определяя, что первый сегмент (128) и первый сегмент подструктуры (620) имеют общий пиксел и что второй сегмент (142) и вторая подструктура (636) имеют общий пиксел, причем компьютер определяет, что третий сегмент (150) не имеет общего пикселя с первой и второй подструктурами (620, 636), при этом компьютер присваивает координаты первому и второму сегментам (128, 142) таким образом, что первый и второй сегменты ближе друг к другу в перцептивном пространстве, чем третий сегмент (150) к первому и второму сегментам (128, 142).
5. Способ по пункту 2, в котором способ дополнительно содержит этапы идентификации компьютером объектов при первом масштабе увеличения одновременно с идентификацией подструктур при втором масштабе увеличения.

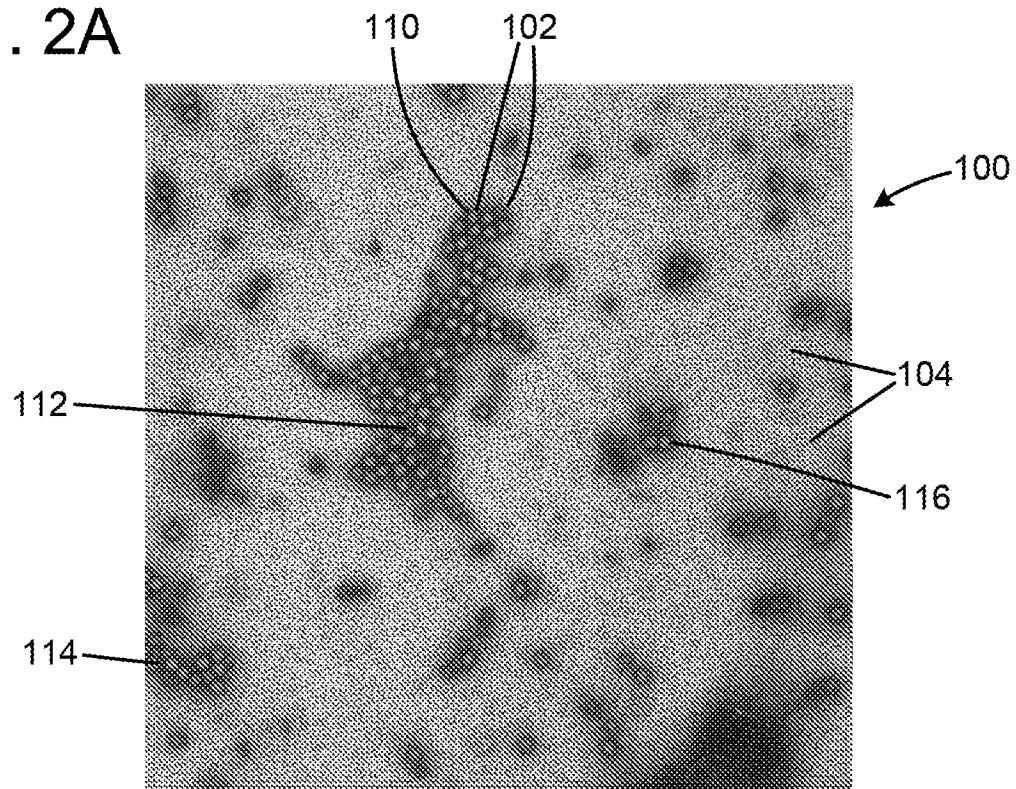
6. Способ по пункту 3, в котором способ дополнительно содержит этапы, на которых идентифицируют третий объект (150) на втором цифровом изображении.
- 5 7. Способ по пункту 1, в котором способ дополнительно содержит этап, на котором идентифицируют четвертый объект (222), к которому прикреплены подструктуры (230, 232, 234, 236, 238, 240), преобразуют объект (222) и подструктуры (230-240) в объект (222') и подструктуры (230'-240'), при сравнении неравенства треугольников анализируют как подструктуры (230'-240') соотносятся
- 10 друг с другом по сравнению с тем, как подструктуры (130'-140') соотносятся друг с другом и как подструктуры (144', 146' и 148') соотносятся друг с другом, для определения сходств.

Фиг. 1

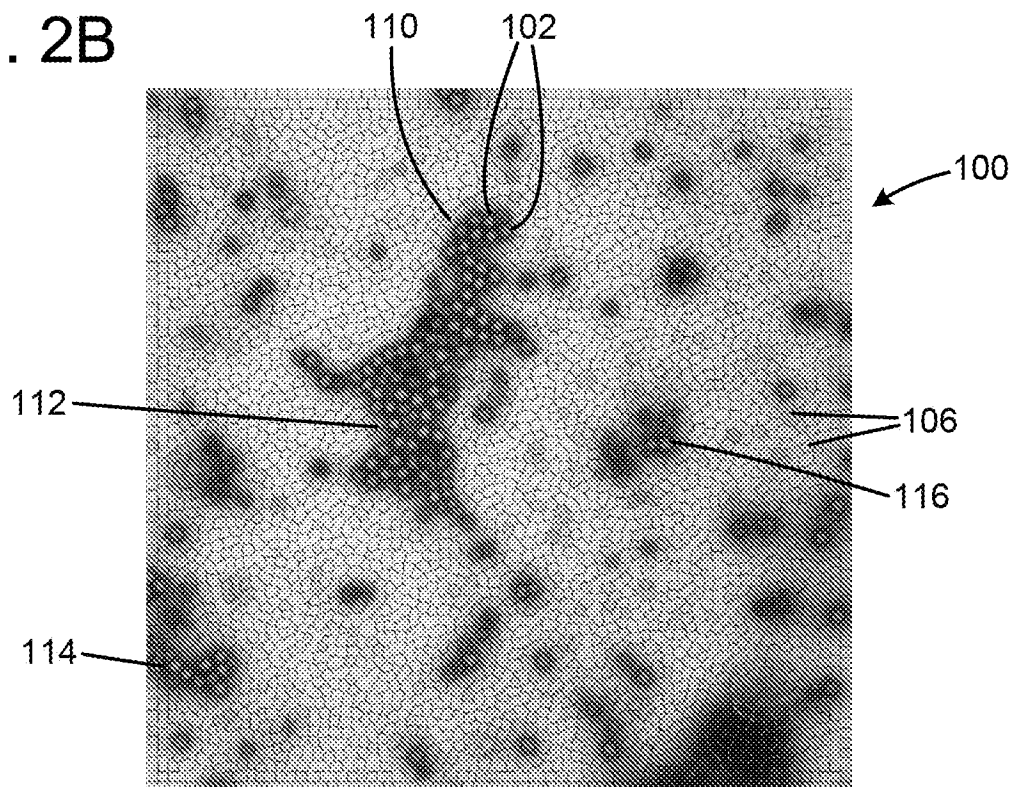


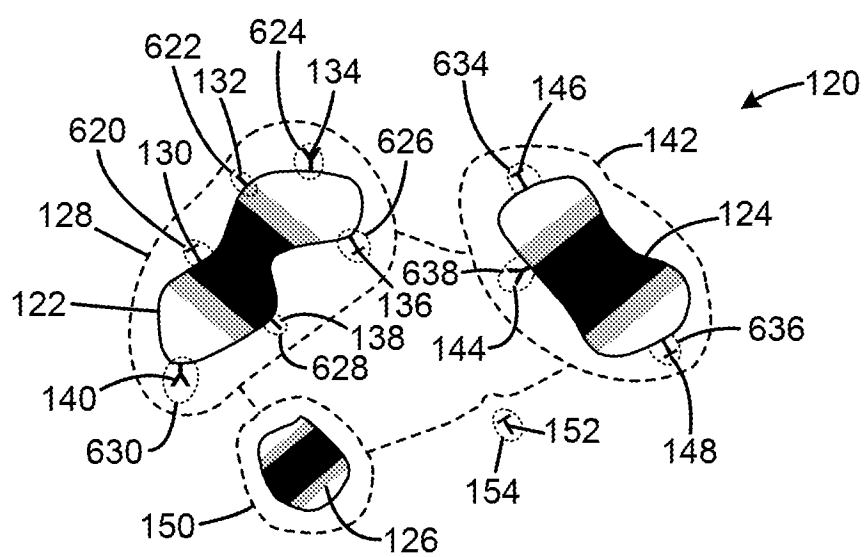
2/10

Фиг. 2А

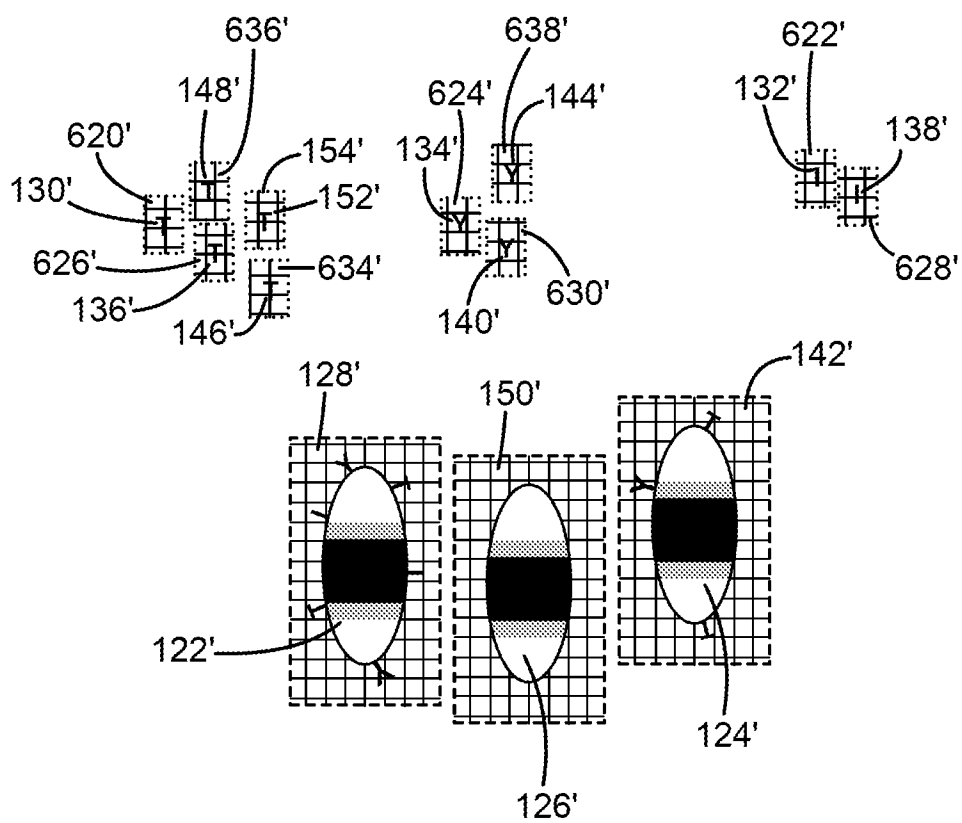


Фиг. 2В

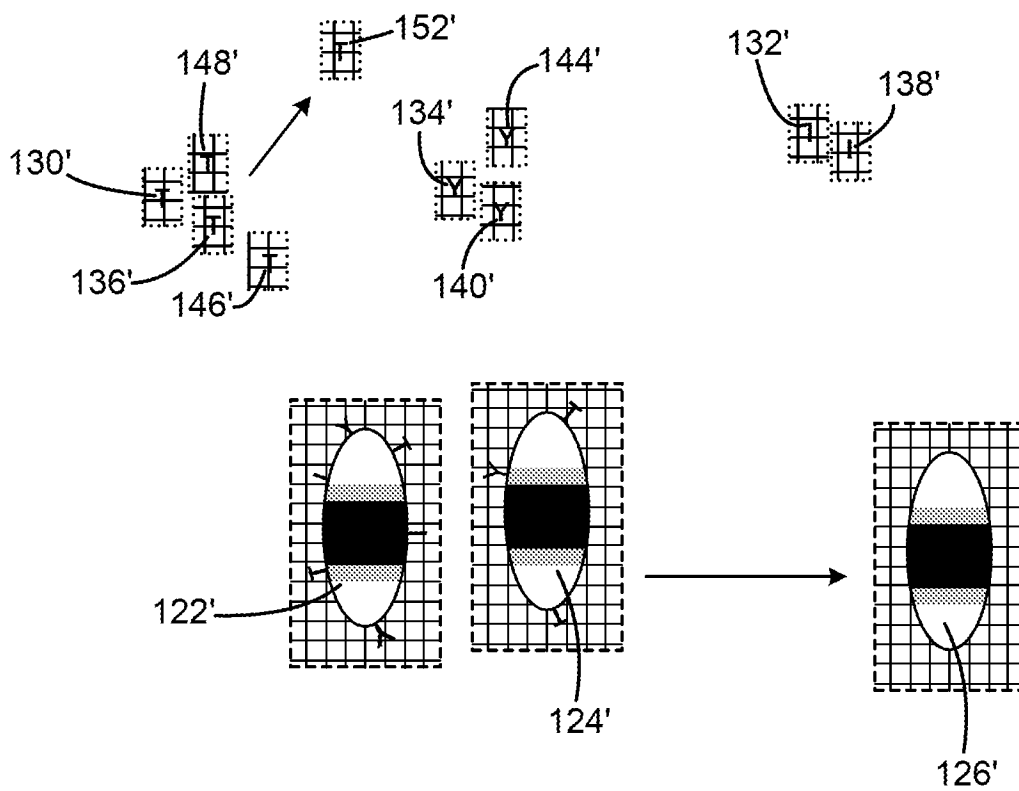




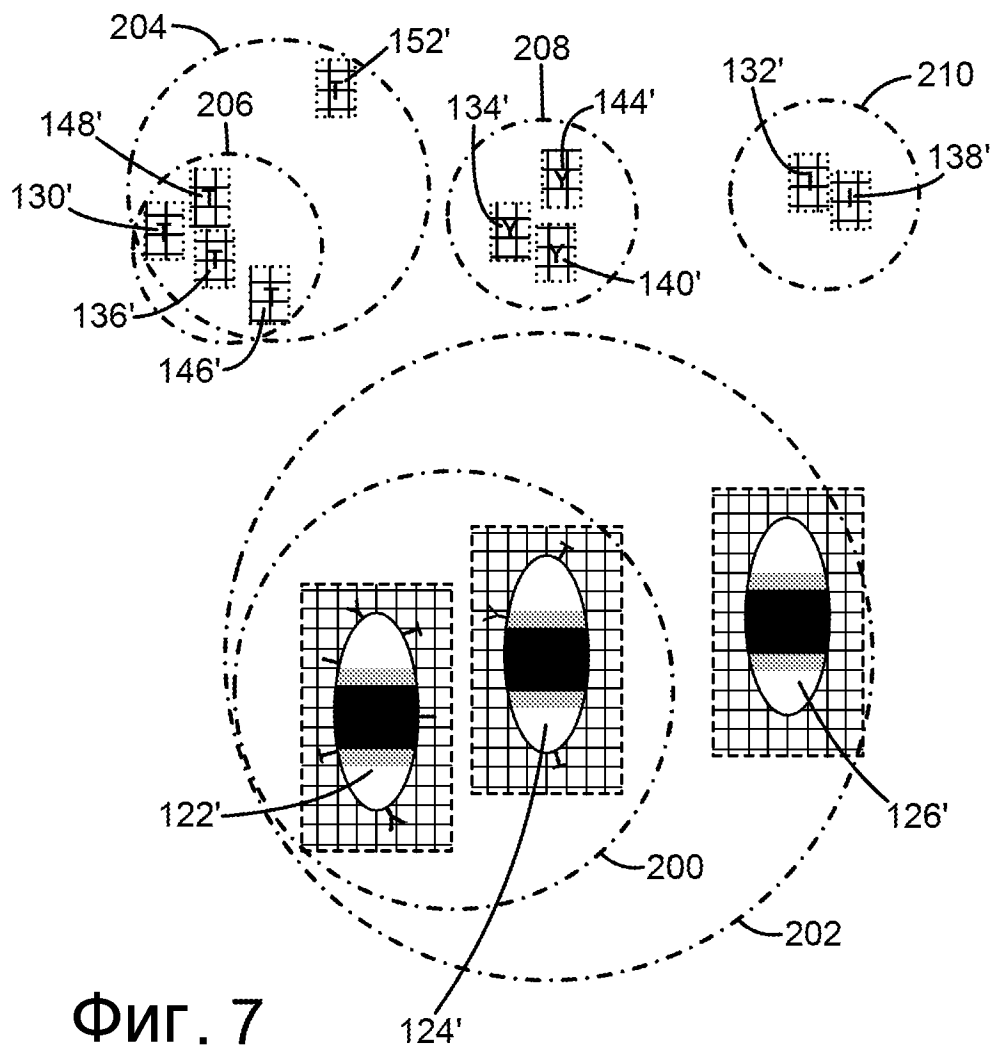
ФИГ. 3



Фиг. 5



Фиг. 6

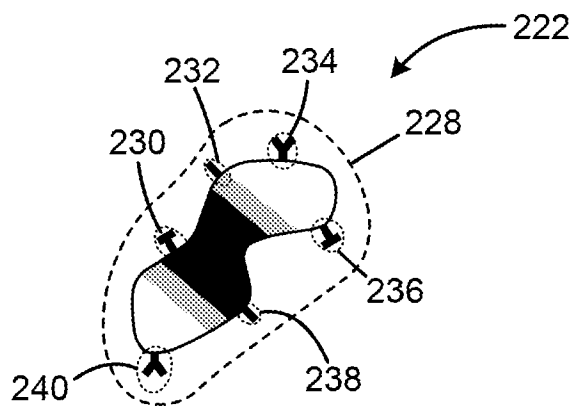


Фиг. 7

8/10

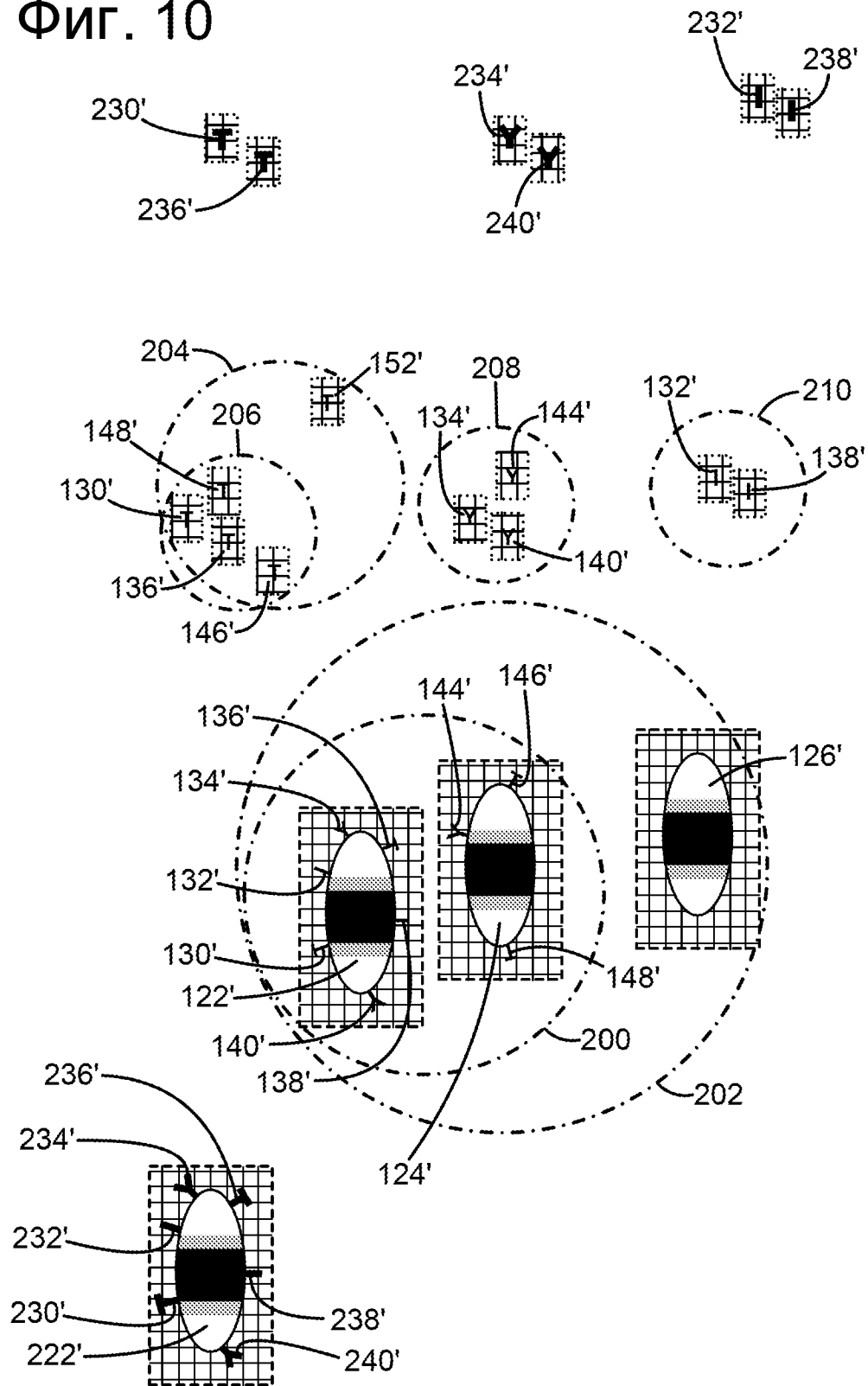


Фиг. 8



Фиг. 9

Фиг. 10



Фиг. 11

