

Проведено аналіз меж застосуванням різноманітних статистичних моделей вихідних сигналів в оптико-електронних системах. Подано залежність граничних законів розподілення флуктуацій вихідного сигналу від коефіцієнту послаблення нейтрального фільтру, що включено до складу оптичної ланки системи. Надано порівняльний аналіз помилок виявлення сигналу у залежності від обраної статистичної моделі вихідних сигналів, що базуються на центральній та узагальнених граничних теоремах

Ключеві слова: оптико-електронні системи, Пуассонівська модель сигналів, флуктуації сигналу, гаусова та негаусові статистики

Проведен анализ границ применимости различных статистических моделей выходных сигналов в оптико-электронных системах. Представлена зависимость предельных законов распределения флуктуаций выходного сигнала от коэффициента ослабления нейтрального фильтра, включенного в состав оптического звена системы. Представлен сравнительный анализ ошибок обнаружения сигнала в зависимости от выбранной статистической модели выходных сигналов на основе центральной и обобщенных предельных теорем

Ключевые слова: оптико-электронные системы, Пуассоновская модель сигналов, флуктуации сигнала, гауссова и негауссовы статистики

УДК 621.37:621.391

DOI: 10.15587/1729-4061.2015.39950

ИСПОЛЬЗОВАНИЕ УСТОЙЧИВЫХ ЗАКОНОВ РАСПРЕДЕЛЕНИЯ ПРИ ОЦЕНКЕ ЭФФЕКТИВНОСТИ ОБРАБОТКИ СИГНАЛОВ В ОПТИКО- ЭЛЕКТРОННЫХ СИСТЕМАХ

Т. А. Стрелкова

Кандидат технических наук, доцент

Кафедра основ радиотехники

Харьковский национальный

университет радиоэлектроники

пр. Ленина, 14, г. Харьков, Украина, 61166

E-mail: strelkova_t@ukr.net

1. Введение

Описание сигналов в оптико-электронных системах основано на различных физико-математических и статистических моделях, учитывающих как параметры среды распространения, так и характеристики сигналов и помех. Теоретические положения волновой теории света, как правило, используются при приеме и обработке оптических сигналов больших интенсивностей. При приеме слабых сигналов (например, сигналов от слабых космических объектов, сигналов в условиях низкой освещенности) применяется корпускулярная теория света, дающая возможность учесть квантовую структуру оптического излучения.

Разрабатывая оптимальные и квазиоптимальные алгоритмы обнаружения сигналов на фоне помех, необходимо учитывать вероятностный характер, как сигнальной, так и фоновой компонент. Вероятностный характер выходных сигналов определяется факторами, влиянием которых нельзя пренебрегать. Вклад во флуктуации сигнала вносят, например, процессы, происходящие в трассе распространения оптического излучения; в фотоприемном устройстве; корпускулярные свойства оптического излучения. Также входное оптическое излучение, взаимодействуя с элементами системы, может претерпевать пространственно-координатные и пространственно-спектральные из-

менения, которые оказывают влияние на статистические характеристики выходных сигналов оптико-электронных систем.

Эффективность обработки сигналов и качество алгоритмов обнаружения в оптико-электронных системах оценивается с помощью условных вероятностей принятия решения. Расчет условных вероятностей ошибок первого и второго рода проводят на основе принятой (выбранной) статистической модели выходных сигналов и помех.

Исследование процессов взаимодействия входного оптического излучения с элементами системы и учет дополнительных факторов, оказывающих влияние на статистические свойства выходных сигналов – закономерный процесс развития оптических и оптико-электронных систем.

Изменение статистических свойств выходных сигналов, и, как следствие, выбор статистической модели, может повлиять на точность определения параметров выходных сигналов и не позволит обеспечить потенциальную точность при получении результатов с наилучшими вероятностными характеристиками о наблюдаемых объектах и их параметрах. Определение плотности вероятности и предельных законов распределения выходных сигналов является на современном этапе развития оптико-электронных систем основной задачей при оценке потенциальных возможностей систем.

2. Анализ литературных данных и постановка задачи

На рис. 1 представлено дерево проблем приема и обработки оптических сигналов в оптико-электронных системах.

Учитывая различные условия работы оптико-электронных систем и свойства оптического излучения, проведем анализ границ применимости различных статистических моделей выходных сигналов, лежащих в основе разработки алгоритмов обработки сигналов.

На формирование статистического поведения выходных сигналов оптико-электронных систем оказывают влияние свойства оптического излучения, свойства среды распространения, свойства оптического звена и свойства приемника излучения.

Составляя математические модели принимаемого оптического излучения, используются нормальное, логарифмически-нормальное, экспоненциальное распределения [1–3]. Большое распространение получили математические модели принимаемых сигналов, составленные на основе пуассоновской статистики [1, 4–7]. Эти модели учитывают корпускулярную структуру потока фотонов, условия формирования оптического излучения, отсутствие взаимодействия между фотонами в потоке, а также стохастические свойства оптических сигналов при анализе квантовых шумов. Однако при составлении математической модели сигналов на основе пуассоновской статистики, как правило, делаются некоторые допущения и принимаются ограничения свойств оптических сигналов, сводящие их к трем основным свойствам простейшего (пуассоновского) потока – стационарности, ординарности и отсутствия последствия. Это позволяет ис-

пользовать свойство инвариантности пуассоновского потока при описании его взаимодействия с элементами системы. Применение такого математического аппарата позволяет получить аналитические выражения для основных показателей эффективности системы и создать высокоэффективные алгоритмы обработки сигналов в оптико-электронных системах [8–10].

При оценке качества алгоритмов обнаружения случайных сигналов на фоне случайных помех используются предельные распределения, характеризующие выходной сигнал. Выходной сигнал оптико-электронных систем представляется как сумма большого числа независимых случайных величин, которая в соответствии с центральной предельной теоремой стремится к нормальному закону распределения как предельному.

В процессе регистрации оптического излучения во многих эмпирических исследованиях флуктуации принимаемого оптического излучения отличаются от принятых моделей. Исследования показывают изменение дисперсии принимаемых сигналов относительно среднего значения, и используется субпуассоновская статистика [11–13]. Отличия зачастую связывают со свойствами среды распространения оптических сигналов; с характеристиками элементов системы, например, рассеивающие свойства, дефекты и загрязнение оптических элементов; ограниченной динамическим диапазоном фотоприемника; конечностью времени регистрации, а также с квантовой природой оптического излучения. Однако для оценки качества алгоритмов обработки сигналов также используют нормальный закон распределения как предельный.



Рис. 1. Дерево проблем приема и обработки оптических сигналов в оптико-электронных системах

В последнее время большой интерес вызывают исследования стохастических процессов, которые не подчиняются гауссовой статистике и классической центральной предельной теореме. Статистической особенностью таких процессов является существенно более высокая вероятность возникновения больших флуктуаций. В таких случаях используются обобщенные предельные теоремы. Распределение суммы независимых случайных величин описывается семейством устойчивых распределений.

Понятие устойчивых распределений введено в 1925 году П. Леви в результате изучения свойств сумм одинаково распределенных случайных величин. Этот класс распределений включает в себя распределения с тяжелыми хвостами и несимметричными плотностями распределения. Часто при анализе статистических зависимостей пренебрегают возможностью крупных событий, лежащих на «хвосте» распределения. Распределения с тяжелыми хвостами – распределения, хвост которых нельзя «отрезать», то есть нельзя пренебречь влиянием крупных, но редких событий. Под крупными событиями понимают явления, влияние от которых может превосходить ущерб от всех остальных событий этого класса.

Негауссовы случайные процессы используются при описании транспорта зарядов на поверхности полупроводника, взаимодействия оптического излучения со стеклом [14–20]. Негауссовы изотропные распределения обладают несимметричными свойствами, дисперсия потока, α , следовательно, и все высшие моменты устойчивых распределений бесконечны – это связано с поведением хвостов плотностей вероятностей при больших значениях случайных величин. Распределение временных характеристик наступления событий в потоке имеет не экспоненциальный, а обратно-степенной вид.

При использовании центральной предельной теоремы элементарные отклонения, вызывающие отклонение суммы независимых случайных величин (флуктуации сигнал и шума), принимаются сравнимыми по порядку их влияния на рассеивание суммы. Вероятность появления значений больших интенсивностей пренебрежимо мала. Использование обобщенных предельных теорем позволяет описать стохастическое поведение случайной величины с учетом того факта, что вероятность регистрации больших значений интенсивности не является пренебрежимо малой. Элементарные отклонения, влияющие на сумму случайных величин, могут иметь иерархическую структуру. На формирование стохастического процесса – выходного сигнала опико-электронных систем, могут оказывать влияния факторы, которые принципиально не устранимы и при описании статистических свойств могут использоваться различные предельные законы.

При установлении границ применимости теоретических моделей основную роль играет интенсивность принимаемых сигналов по отношению к интенсивности шумовой составляющей. Работая на границах динамического диапазона систем, т. е. на «хвостах» плотности распределения выходных сигналов, необходимо учитывать, что вклад больших флуктуаций будет существенным при оценке обнаружительной способности системы.

3. Цель и задачи исследования

Целью работы является учет изменений статистического поведения выходных сигналов опико-электронных систем при выборе математической модели выходных сигналов и оценке обнаружительной способности системы.

Для достижения поставленной цели были поставлены следующие задачи:

– провести анализ изменений статистических свойств выходных сигналов опико-электронных систем с ограниченным динамическим диапазоном при формировании пространственно-временного распределения оптического излучения в плоскости фотоприемника;

– учесть асимптотическое поведение «хвостов» плотностей распределения выходных сигналов опико-электронных систем при использовании пороговых методов обнаружения;

– провести расчет условной вероятности ложной тревоги системы на основе различных статистических моделей выходных сигналов.

4. Изменение статистических свойств выходных сигналов опико-электронных систем от состава и качества оптического звена

Одним из важнейших элементов опико-электронных систем является оптическое звено, основной функцией которого является построение изображения. Сегодня достигнуты существенные результаты в создании высококачественных оптических приборов нового поколения, разработаны и реализованы новые марки оптических стекол, к которым предъявляются весьма строгие требования. Состав оптического звена изменяется в зависимости от условий работы системы. При регистрации сигналов больших интенсивностей в составе оптического звена используют нейтральные фильтры для согласования динамического диапазона фоточувствительных элементов с интенсивностью принимаемого сигнала.

При описании процесса взаимодействия оптического излучения с оптическим звеном, оптические стекла в основном рассматриваются с точки зрения макроскопической теории упорядоченности и симметрии. Однако, описывая взаимодействие квантов света с материалом оптического стекла, необходимо учесть, что в таких материалах может возникать пространственная разупорядоченность, вызванная наличием хаотически расположенных дефектов, возникающих вследствие, влияния процессов атмосферных воздействий и старения, а также стохастического расположения центров окраски в оптически плотных материалах (например, нейтральных фильтрах). Неупорядоченные материалы на микроскопическом уровне принципиально неоднородны и могут различаться по структуре. В различных участках макроскопически однородного оптического стекла, процессы трансформации энергии фотонов могут вносить дополнительные флуктуации.

В нелинейных статистически неоднородных средах положение и ориентация отдельных частиц (дефектов) можно считать неизвестными. Описывая взаимодействие оптического излучения с неупорядоченными

структурами (например, оптическим звеном), можно представить, что фотоны входят в аморфный материал, рассеиваются, отражаются, поглощаются, активируют колебательные процессы (фотон-фононное взаимодействие) внутри и выходят с противоположной стороны, описание процесса прохождения излучения внутри материала является достаточно сложным процессом (рис. 2).

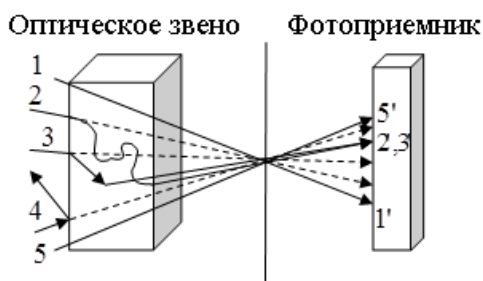


Рис. 2. Формирование пространственного распределения оптического излучения в плоскости фотоприемника

Такой процесс можно описать как случайное блуждание, то есть, фотон, проходя расстояние в аморфном веществе, может потерять часть своей энергии во время поглощения, изменить свое направление движения при взаимодействии с дефектами стекла. Фотон, проходя через решетку аморфного вещества, может возбудить в ней фонон, растратив при этом часть своей энергии, вследствие чего его частота уменьшается, и возникнут кванты другой энергии. Если в веществе уже был возбужденный фонон, пролетающий фотон может поглотить его, увеличивая, за этот счет свою энергию и возникнут фотоны с большей энергией. Поэтому, когда фотоны выходят из материала, то оптическое излучение будет иметь меньшую интенсивность и в общем случае характеризоваться другим спектральным составом. Атмосферное воздействие и процессы старения стекла могут изменить коэффициенты рассеяния и отражения это в свою очередь может привести к тому, что свет входит в материал, многократно отражается внутри, изменяет свою траекторию движения и выходит наружу не в ожидаемом месте, нарушая при этом принцип линейной суперпозиции.

Таким образом, процесс взаимодействия фотонов с оптическим звеном системы можно рандомизировать не только по уровню интенсивности и частотному составу входного и выходного сигнала, но и по траектории прохождения через оптическое звено, и как следствие пространственно-временному распределению оптического излучения в плоскости фотоприемника.

В [20, 21] представлен отчет о проведенных экспериментальных исследованиях процесса ослабления оптических потоков в оптико-электронных системах. Исследования включают изучение статистических характеристик выходных сигналов оптико-электронных систем при использовании различных нейтральных фильтров для ослабления входных оптических потоков. Представлена зависимость предельного закона распределения флуктуаций выходного сигнала от коэффициента ослабления нейтрального фильтра k , включенного в состав оптического звена системы (рис. 3).

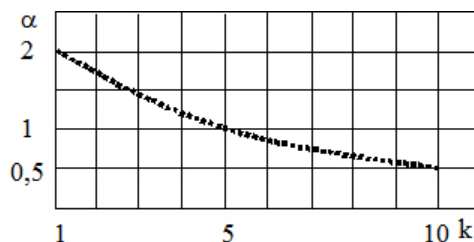


Рис. 3. Экспериментальная зависимость коэффициента ослабления оптического звена k и характеристического показателя $0 < \alpha < 2$

Исследования асимптотического поведения хвостов плотностей распределения выходных сигналов оптико-электронных систем показали возможность использования устойчивых законов распределения для описания выходных сигналов оптико-электронных систем, а также установить зависимость коэффициента ослабления оптического звена k и характеристического показателя $0 < \alpha < 2$.

5. Оценка обнаружительной способности оптико-электронных систем

Традиционные методы обнаружения сигналов в оптико-электронных системах основаны на пороговой обработке сигнала, то есть на сравнении величины отклика фоточувствительных элементов на воздействие аддитивной смеси сигнальной и фоновой составляющей с установленным пороговым значением, величина которого обусловлена выбранным критерием качества принятия решения. Решение об обнаружении сигнала принимается, когда амплитуда электрического сигнала, сформированного одним элементом разрешения матрицы фотоприемников, превышает установленное значение порога по различным критериям, например критерию Неймана-Пирсона [1, 2, 9]. В этом случае решающее правило будет состоять в вычислении значения отношения правдоподобия и сравнения результата вычислений с пороговым значением. Выражение для отношения правдоподобия, при разбиении интервала наблюдения i на статистически независимые подинтервалы, имеет вид:

$$\Lambda = \prod_{i=1}^m \frac{P_{n_i}^{(cn)}}{P_{n_i}^{(fn)}} \tag{1}$$

Плотность вероятности сигнальной и фоновой компонент, используя общепринятую пуассоновскую статистику можно определить соответственно:

$$P_{n_i}^{(cn)} = \frac{(\bar{n}_{ci} + \bar{n}_{mi}/k)^{n_i}}{n_i!} \cdot e^{-\left(\frac{\bar{n}_{ci} + \bar{n}_{mi}}{k}\right)}, \tag{2}$$

$$P_{n_i}^{(fn)} = \frac{(\bar{n}_{mi}/k)^{n_i}}{n_i!} \cdot e^{-\left(\frac{\bar{n}_{mi}}{k}\right)}, \tag{3}$$

где k – коэффициент ослабления нейтрального фильтра; \bar{n}_{ci} , \bar{n}_{mi} – средние значения сигнальной и фоновой компонент на i -м интервале наблюдения.

Соответственно логарифм отношения правдоподобия запишем:

$$\ln \Lambda = \ln \prod_i \frac{P_{n_i}^{(cn)}(t_i)}{P_{n_i}^{(n)}(t_i)} = - \sum_{i=1}^m \bar{n}_{ci} k^{-1} + \sum_{i=1}^m \ln \left(\frac{((\bar{n}_{ci} + \bar{n}_{ni})k^{-1})^{n_i} / n_i!}{(\bar{n}_{ni} k^{-1})^{n_i} / n_i!} \right). \tag{4}$$

Качество алгоритмов обработки и обнаружения сигналов определяется значениями вероятностных характеристик, таких как условные вероятности правильного обнаружения и ложной тревоги F (выражения (5) и (6) соответственно) используя центральную предельную теорему:

$$D = 0,5 - \frac{1}{\sqrt{2\pi}} \int_{n_0}^{\infty} e^{-\frac{(n_c + n_n)k^{-1}}{2}} d(n_c + n_n), \tag{5}$$

$$F = 0,5 - \frac{1}{\sqrt{2\pi}} \int_{n_0}^{\infty} e^{-\frac{(n_n k^{-1})^2}{2}} d(n_n), \tag{6}$$

где n_0 – порог обнаружения.

Принимая во внимание возможность применения различных предельных распределений выходных сигналов и учитывая факторы и условия, при которых нет возможности использовать известную предельную статистику и, следовательно, алгоритмы обработки сигналов, разработанные на ее основе, представляется целесообразным проанализировать ошибки обнаружения сигналов в оптико-электронных системах с использованием устойчивых законов распределения.

В общем случае формирование выходного сигнала оптико-электронных систем представляется как стохастический процесс, используя корпускулярную теорию, потоком дискретных частиц. Как обсуждалось ранее в основе формирования такого стохастического процесса лежит Пуассоновский закон. Функция является разрывной и принимает только целочисленные значения, то есть изменяется резкими скачками. Число реализаций события независимы в течение непрерывающихся интервалов времени и для любого малого интервала времени вероятность осуществления события очень мала. Такая функция характеризуется подвижными разрывами непрерывности, места которых заранее неизвестны. На основе пуассоновских процессов в [19] построен широкий класс процессов, позволяющих, используя уравнение аддитивных процессов, описать выходной сигнал системы.

Аддитивный процесс определяется:

$$X(t) = f(t) + X_0(t) + X_1(t) + X_2(t), \tag{7}$$

где $f(t)$ – неслучайная функция; $X_0(t), X_1(t), X_2(t)$ – случайные члены суммы стохастически независимые друг от друга, определяющие функцию фиксированных разрывов, функцию приводящую к броуновскому движению и функцию суммы подвижных разрывов.

Для вероятностных распределений таких процессов не выполняется центральная предельная те-

орема. Увеличение выборки, каждый раз конечной, не повышает точность, а снижает ее. Функция распределения:

$$F(t) \sim C_1 \frac{H_1(x)}{|x|^\alpha} \quad 1 - F(t) \sim C_2 \frac{H_2(x)}{x^\epsilon},$$

$$C_1 \geq 0, C_2 \geq 0, C_1 + C_2 > 0, 0 < \alpha < 2.$$

Предельный закон распределения выходного сигнала в зависимости от коэффициента нейтрального фильтра k можно определить согласно рис. 3:

$$P_{n_i}^{(n)} = \frac{1}{2\sqrt{\pi}} \exp\left(-\frac{(n_{ni})^2}{4}\right) \text{ при } k \approx 1; \tag{8}$$

$$P_{n_i}^{(n)} = \left[\pi \left(1 + (n_{ni})^2\right)\right]^{-1} \text{ при } k \approx 5; \tag{9}$$

$$P_{n_i}^{(n)} = \frac{1}{2\sqrt{\pi}} (n_{ni})^{-3/2} \exp\left(-\frac{1}{4n_{ni}}\right) \text{ при } k \approx 10. \tag{10}$$

Изучение стохастического поведения входных сигналов оптико-электронных систем в различных условиях и использование обобщенных предельных теорем показывает, что при выборе порога обнаружения n_0 на уровне, например 3σ , где σ – среднее квадратическое отклонение для гауссовой статистики, вероятность ошибочного решения будет существенно отличаться от ошибочного решения для статистики устойчивых законов (рис. 4).

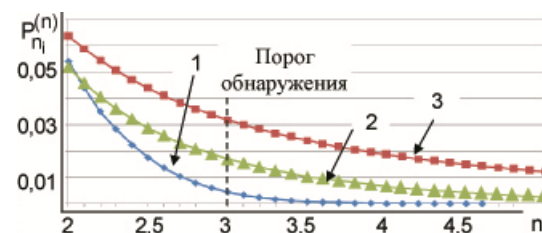


Рис. 4. Выбор порога обнаружения сигналов в оптико-электронных системах: 1 – нормальный закон распределения; 2 – устойчивый закон распределения $\alpha=0,5$; 3 – устойчивый закон распределения $\alpha=1$

Проведем расчеты условной вероятности ложной тревоги F (рис. 5), согласно выражению (6), используя предельные распределения (выражения (8), (9), (10)), при различных порогах обнаружения $n_0 = \sigma$, где $\sigma = 2, \dots, 7$ для гауссовой статистики.

Расчеты показали, что при малых коэффициентах ослабления оптического звена системы, наблюдается удовлетворительное совпадение значений вероятности ложной тревоги при выборе порогового значения до $n_0 = 3\sigma$ для различных предельных законов распределений. При увеличении коэффициента ослабления, асимптотическое поведение хвоста плотности распределения изменяется, что приводит к существенному увеличению вероятности ложной тревоги.

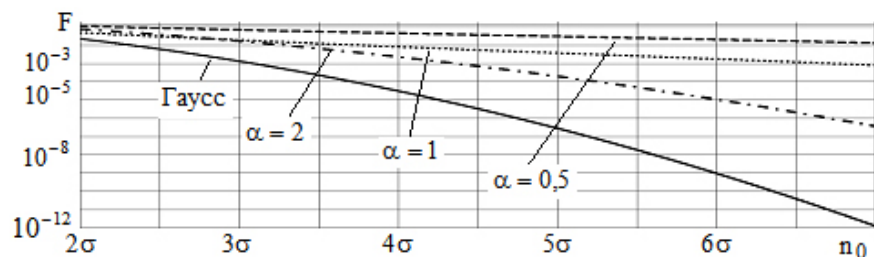


Рис. 5. Условная вероятность ложной тревоги при различных порогах обнаружения $n_0 = \sigma$ для гауссовой статистики

6. Выводы

Основным этапом при разработке алгоритмов обработки сигналов и оценке обнаружительной способности оптико-электронных систем является выбор физико-математической и статистической модели выходных сигналов. Условия работы системы и пространственно-энергетические характеристики принимаемого излучения являются существенными при установлении границ применимости статистических моделей на основе центральной и обобщенных предельных теорем. Проведенный анализ статистических характеристик

выходных сигналов оптико-электронных систем с ограниченным динамическим диапазоном, в составе которых используются оптически плотные материалы, показывает удовлетворительное согласование статистических моделей на основе центральной предельной теоремы с экспериментальными данными в диапазоне до 3 СКО. Однако асимптотическое поведение функции распределения выходных сигналов в оптико-электронных системах с ограниченным динамическим диапазоном варьируется в зависимости от состава и качества оптического звена системы. Как показали расчеты вероятности ложной тревоги системы, использование статистических моделей на основе центральной предельной теоремы при обнаружении сигналов, дает значения условной вероятности ложной тревоги, заниженное по сравнению со статистическими моделями на основе обобщенных предельных теорем. Совершенствование алгоритмов обработки сигналов с учетом статистической модели выходных сигналов на основе устойчивых законов позволит избежать ошибок обнаружения.

Литература

1. Оптическая локация. Теоретические основы приема и обработки оптических сигналов [Текст]: монография / под ред. А. И. Стрелкова. – Х.: Апостроф, 2010. – 312 с.
2. Мосягин, Г. М. Теория оптико-электронных систем [Текст]: монография / Г. М. Мосягин, В. Б. Немтинов, Е. Н. Лебедев. – Москва, 1990. – 432 с.
3. Кокс, Д. Теория восстановления [Текст]: монография / Д. Кокс, В. Смит. – Москва, 1967. – 300 с.
4. Лытюга, А. П. Математическая модель сигналов в телевизионных системах при наблюдении низкоорбитальных космических объектов в дневное время [Текст] / А. П. Лытюга // Зб. наук. праць Харківського університету Повітряних Сил. – 2009. – Вип. 4 (22). – С. 41–46.
5. Федосеев, В. И. Прием пространственно-временных сигналов в оптико-электронных системах (пуассоновская модель) [Текст]: монография / В. И. Федосеев. – Сп-б.: Университетская книга, 2011. – 232 с.
6. Гальярди, Р. Оптическая связь [Текст]: монография / Р. Гальярди, Ш. Карп. – М.: Связь, 1978. – 424 с.
7. Yang, F. Bits From Photons. Oversampled Image Acquisition Using Binary Poisson Statistics [Text] / F. Yang, Y. M. Lu, L. Sbaiz, M. Vetterli // IEEE Transactions on image processing. – 2012. – Vol. 21, Issue 4. – P. 1421–1436. doi: 10.1109/tip.2011.2179306
8. Березин, В. В. Обнаружение и оценивание координат изображений точечных объектов в задачах астронавигации и адаптивной оптики [Текст] / В. В. Березин, А. К. Цыбулин // Вестник ТОГУ. – 2008. – № 1(8). – С. 11–20.
9. Никитин, В. М. Адаптивная помехозащита оптико-электронных информационных систем [Текст]: монография / В. М. Никитин, В. Н. Фомин, А. И. Николаев, И. Л. Борисенков. – Белгород, 2008. – 196 с.
10. Strelkov, A. Signal Detection in Technical Vision Systems [Text] / A. Strelkov, Ye. Zhilin, A. Lytyuga, S. Lisovenko // Telecommunications and Radio Engineering. – 2007. – Vol. 66, Issue 4. – P. 283–293. doi: 10.1615/telecomradeng.v66.i4.10
11. Большаков, И. А. Прикладная теория случайных потоков [Текст]: монография / И. А. Большаков, В. С. Ракошиц. – Москва, 1978. – 248 с.
12. Белоусов, Ю. И. Учет характеристик флуктуаций фонового излучения пригоризонтной области моря в алгоритмах обработки сигналов инфракрасных приборов [Текст] / Ю. И. Белоусов, Д. В. Иванов // Изв. Вузов. Приборостроение. – 2008. – Т. 52, № 8. – С. 43–49.
13. Глаубер, Р. Оптическая когерентности и статистика фотонов [Текст]: монография / Р. Глаубер. – М.: Наука, 1966. – 452 с.
14. Шелухин, О. И. Негауссовские процессы в радиотехнике [Текст]: монография / О. И. Шелухин. – М.: Радио и связь, 1998. – 310 с.
15. Сибатов, Р. Т. Дробко-дифференциальная кинетика переноса заряда в неупорядоченных полупроводниках [Текст] / Р. Т. Сибатов, В. В. Учайкин // Физика и техника полупроводников. – 2007. – Т. 41, Вып. 3. – С. 346–351.
16. Harvey, J. E. Image degradation due to scattering effects in two-mirror telescopes [Text] / J. E. Harvey, N. Choi, A. Krywonos, G. Peterson, M. Bruner // Optical Engineering. – 2010. – Vol. 49, Issue 6. – P. 063202. doi: 10.1117/1.3454382
17. Sabathil, M. Opto-electronic and quantum transport properties of semiconductor nanostructures [Text] / M. Sabathil // Vol. 67 of Selected Topics of Semiconductor Physics and Technology (Verein zur Förderung des Walter Schottky Instituts, Garching, 2004).

18. Золотарев, В. М. Устойчивые законы и их применение [Текст]: монография / В. М. Золотарев. – Москва, 1984. – 66 с.
19. Леви, П. Стохастические процессы и броуновское движение [Текст]: монография / П. Леви. – М.: Наука, 1972. – 375 с.
20. Стрелкова, Т. А. Статистические свойства выходных сигналов оптико-телевизионных систем с ограниченным динамическим диапазоном [Текст] / Т. А. Стрелкова // Восточно-Европейский журнал передовых технологий. – 2014. – Т. 2, № 9 (68). – С. 38–44. doi: 10.15587/1729-4061.2014.23361
21. Strelkova, T. A. Studies on the Optical Fluxes Attenuation Process in Optical-electronic Systems [Text] / T. A. Strelkova, // Semiconductor physics, quantum electronics & optoelectronics (SPQEO). – 2014. – Vol. 17, Issue 4. – P. 421–424.

В даній статті представлений спосіб автоматизованої оцінки числа кластерів, заснований на евристичному підході чіткої кластеризації вхідного масиву даних з використанням густини розподілу даного масиву. Спосіб включає в себе правило прийняття рішення. Даний підхід чіткої кластеризації дає змогу оцінити, чи є число кластерів більше за одиницю

Ключові слова: спосіб кластеризації даних, кластер, евристичний підхід, густина розподілу

В данной статье представлен способ автоматизированной оценки числа кластеров, основанный на эвристическом подходе четкой кластеризации входного массива данных с использованием плотности распределения данного массива. Способ включает в себя правило принятия решения. Данный подход четкой кластеризации позволяет оценить, является ли число кластеров больше чем единица

Ключевые слова: способ кластеризации данных, кластер, эвристический подход, плотность распределения

UDC 004.9

DOI: 10.15587/1729-4061.2015.39785

NEW DATA CLUSTERING HEURISTIC ALGORITHM

V. Mosorov

Doctor of Technical Sciences*

E-mail: volodymyr.mosorov@p.lodz.pl

T. Panskyi

Graduate student*

E-mail: panskyi@gmail.com

*Institute of Applied Computer Science

Lodz University of Technology

Stefanowskiego str., 18/22,

Lodz, Poland, 90-924

1. Introduction

The goal of this article is to propose a method of clustering the data by data mining. Clustering is a division of data into groups of similar objects using appropriate algorithms and rules. Clustering has been thoroughly studied and perfected over the years in different areas and fields of science including pattern recognition [1], machine learning [2], statistics and image processing [3]. Data clustering has been applied in different fields of science and industry, such as biology and bioinformatics, medicine, marketing, computer science, social science, robotics, mathematical chemistry, climatology, physical geography and others.

Traditional clustering methods and algorithms are computationally expensive when clustering is applied to large data sets. There are three possible cases when the data set can be determined as large: when in one data set there are a lot of elements, when one element has many special properties and when the correct definition of a large number of clusters is problematic. In our case the large data set is the set with a lot of elements [4]. Considering the development of computer technology, consumption of time and computer memory to perform clustering is not taken into account.

2. Analysis of published data and problem statement

Application of various clustering algorithms is impressive: from the clustering of textual data, images, videos to a net-

work clustering algorithm along with the real-time streaming data and outlier detection. Annually old algorithms are modernized and improved both a new ways and approaches of clustering techniques are investigated. Clustering algorithms can be broadly classified [5] into three categories: partitioning, hierarchical and density based. The proposed algorithm has found his niche in density based algorithms as a basic input parameter is the density distribution of a primary data. This algorithm has certain advantages in comparison with existing density based algorithms that will be discussed after viewing the publications. In comparative analysis, attention was paid to some representatives of density based clustering algorithms: DBSCAN [6], FDBSCAN [7], ODBSCAN [8], VDBSCAN [9], ST-DBSCAN [10], Incremental DBSCAN [11], and RDBC [12]. Despite the advantages such as: an opportunity to find arbitrary shape cluster, remove noise from the dataset, cluster spatial-temporal data according to non-spatial, these algorithms has the main drawback – determining the initial parameters for the correct application of the algorithm, such as radius, minpts, number of identical circles. Developed algorithm does not use any of the above input parameters only the initial parameter is the input data set.

3. Purpose and objectives of the study

The key purpose of this paper is determining the quantity of clusters, namely, we are interested in whether there is more than one cluster using only a data set as an input parameter.