

Обнаружение аномалий на пространственно-неоднородных многозональных изображениях

Н.А. Андриянов^а, К.К. Васильев^а, В.Е. Дементьев^а

^а Ульяновский государственный технический университет, 432027 ул. Северный Венец, 32, Ульяновск, Россия

Аннотация

Рассмотрена задача обнаружения аномалий на фоне многомерных изображений. Синтезирован алгоритм, основанный на применении дважды стохастических моделей случайных полей и требующий проведения предварительной фильтрации изображения, например, с помощью векторного фильтра Калмана. Исследована эффективность обнаружения протяженных сигналов на реальных изображениях. Показано, что полученный алгоритм имеет более высокую эффективность, чем известные, основанные на традиционном авторегрессионном описании изображений. Выигрыш объясняется более адекватным описанием неоднородного реального материала дважды стохастическими моделями.

Ключевые слова: дважды стохастические модели; обнаружение аномалий; фильтрация изображений; случайные поля; фильтр Калмана

1. Введение

Задачи обнаружения и оценивания параметров аномалий на изображениях представляют интерес для целого ряда приложений. Среди них можно выделить радио- и гидролокационные комплексы с пространственными антенными решетками, аэрокосмические системы глобального мониторинга Земли, системы технического зрения и др. Для названных систем [1-3] описание сигналов и помех осуществляется с помощью случайных функций нескольких переменных, т. е. многомерных случайных полей (СП). Характерными примерами использования таких СП являются задачи описания и обработки результатов мультиспектральной (до 10 спектральных диапазонов) и гиперспектральной (до 300 диапазонов) съемки участков земной поверхности. При этом необходимо с одной стороны рассматривать аэрокосмические наблюдения как единую многомерную совокупность, а с другой стороны - учитывать ряд характерных особенностей спутниковых изображений, например, выраженную пространственную неоднородность. Среди задач обработки таких изображений особое место занимает задача обнаружения аномалий [4-7]. Такими аномалиями могут быть очаги пожаров, факел стартующей ракеты, полыньи на льду, косяки рыб в океане и пр. При этом фоном для обнаружения являются последовательности многозональных изображений, т.е. снимков территории в разное время в разных спектральных диапазонах. В настоящей работе представлены результаты синтеза и анализа алгоритмов обнаружения аномалий на многозональных спутниковых изображениях.

2. Краткий обзор алгоритмов обнаружения аномалий

Как правило, в задачах обнаружения часто используются статистические алгоритмы обнаружения сигналов (Байеса, Неймана-Пирсона), требующие, однако, достаточного количества априорной информации. Тем не менее, разработка статистических алгоритмов представляет собой актуальную задачу. Во-первых, для подобного рода алгоритмов возможно использование различных математических моделей изображений. Во-вторых, анализ эффективности таких алгоритмов может быть исследован как теоретическим, так и опытным путем. Отличающиеся подходами к определению «аномальности» и базирующиеся на различных моделях изображения алгоритмы [7] были предложены относительно недавно. В их число входят: алгоритм пространственно-спектрального рассогласования, в котором изображение описывается моделью стационарного СП, алгоритм адаптивного спектрального рассогласования, где величина «аномальности» определяется авторами, как ошибка при представлении отсчета через его окрестность, а также вероятностный алгоритм обнаружения аномалий, использующий квантование сигнатур изображений. Стоит отметить, что сравнение работы алгоритмов в работе проводилось только со стандартным RXD-алгоритмом.

Еще одним вариантом является обнаружение аномалий на многомерных сетках с использованием вейвлет-преобразования [5]. Данный метод относится к методам с предварительной обработкой, поэтому с его использованием удастся повысить производительность обнаружения аномалий. Однако затруднительно использование алгоритма с предварительным дискретным вейвлет-преобразованием при решении задач обнаружения аномалий в реальном времени.

В последнее время для обработки гиперспектральных изображений стали применяться топологические инструменты вместе с идеями из теории сетей. Был предложен стандартный алгоритм RX (I. S. Reed, X. Yu), основанный на вычислении стандартных отклонений пикселей от среднего значения в многомерном смысле. Однако хорошо работает он только на простых изображениях, таких как большой лес, но не на сложных городских сценах. Как правило, для сложных изображений используются алгоритмы с переходом к сокращенному описанию, локальные алгоритмы или алгоритмы с предварительной сегментацией [4].

В нашем исследовании мы будем понимать под аномалией априорно заданный и наблюдаемый на многозональном изображении объект. Отметим, что в рамках данной работы параметры сигнала (его значения и местоположение) будем считать известными. В противном случае, потребовалось бы проводить предварительную классификацию аномалий для определения возможных уровней сигнала, а также выполнять поиск таких аномалий не в конкретно заданной области, а по всему изображению.

3. Алгоритмы фильтрации и обнаружения аномалий на фоне дважды стохастических случайных полей

Представим многозональное изображение как совокупность массивов данных, тогда имеем многокомпонентное изображение, состоящее из N компонент, $\{z_{ijk}\}, k=1..N, i=1..M_1, j=1..M_2$, которые получены в результате пространственной дискретизации сигналов, поступивших от различных систем датчиков. При отсутствии полезного сигнала (гипотеза H_0) модель наблюдений можно представить аддитивной смесью:

$$z_{ijk} = x_{ijk} + \theta_{ijk}, (i, j) \in G^k, k=1..N,$$

случайного поля x_{ijk} с нулевым средним и заданной корреляционной функцией (КФ) $B_{(ml)}^k = M\{x_{ij}^k, x_{i+m, j+l}^k\}$ и пространственного белого шума θ_{ijk} с нулевым математическим ожиданием и дисперсией σ_θ^2 в области G , где для всех компонент изображения появление сигнала считается невозможным.

При наличии полезного сигнала (гипотеза H_1) модель наблюдений запишется в виде:

$$z_{ijk} = x_{ijk} + s_{ijk} + \theta_{ijk}, (i, j) \in G_0^k, k=1..N,$$

$$z_{ijk} = x_{ijk} + \theta_{ijk}, (i, j) \notin G_0^k, k=1..N,$$

где G_0^k - область на k -й компоненте изображения, в которой возможно появление полезного сигнала с известными уровнями $s_{ijk}, (i, j) \in G_0^k$. Для упрощения выкладок будем считать, что на каждой из компонент эта область имеет одну и ту же форму: $G_0^k = G_0$. А также в рамках нашей задачи будем считать известной заранее и саму область G_0 .

Общее решение задачи обнаружения основано на построении модифицированного отношения правдоподобия [3]:

$$L = \frac{w(\{z_{ijk}\}/H_1)}{w(\{z_{ijk}\}/H_0)},$$

и сравнении с пороговым значением. По результатам этого сравнения принимается решение в пользу гипотезы о наличии полезного сигнала или гипотезы о его отсутствии.

Исходя из центральной предельной теоремы, выполним аппроксимацию условных плотностей распределения вероятностей $w(\{z_{ijk}\}/H_1)$ и $w(\{z_{ijk}\}/H_0)$ гауссовскими [2,3,8-10]:

$$w(\{z\}/H_1) = \frac{1}{\sqrt{2\pi}\sigma_{z_1}} \exp\left(-\frac{(z - m_{z_1})^2}{2\sigma_{z_1}^2}\right), w(\{z\}/H_0) = \frac{1}{\sqrt{2\pi}\sigma_{z_0}} \exp\left(-\frac{(z - m_{z_0})^2}{2\sigma_{z_0}^2}\right),$$

где m_{z_1} и m_{z_0} - математические ожидания наблюдений $\{z_{ijk}\}$ при наличии полезного сигнала и при его отсутствии соответственно; $\sigma_{z_1}^2$ и $\sigma_{z_0}^2$ - дисперсии наблюдений $\{z_{ijk}\}$ при наличии полезного сигнала и при его отсутствии соответственно.

Тогда оптимальное правило обнаружения сигнала можно записать в виде [3]:

$$L = \bar{s} V_\theta^{-1} (\bar{z} - \hat{x})^T \begin{cases} > L_0 - \text{сигнал есть,} \\ \leq L_0 - \text{сигнала нет,} \end{cases}$$

где V_θ - диагональная матрица со значениями σ_θ^2 , \bar{s} - известный протяженный сигнал, L_0 - порог, который может быть найден исходя из заданной вероятности ложной тревоги.

Для случая отсутствия полезного сигнала оценки \hat{x} являются оптимальными линейными оценками в обычном смысле минимума дисперсии ошибок, основанными на всех имеющихся наблюдениях $\{z_{ijk}\}$. При наличии сигнала полученные величины \hat{x} следует считать псевдооценками, содержащими в своем составе преобразованный входной сигнал \bar{s} .

Таким образом, наилучшая процедура обнаружения предполагает проведение оптимальной фильтрации СП, расчет ковариационной матрицы ошибок фильтрации и выполнение весового суммирования в соответствии с указанными

формулами. Наиболее сложным из указанных этапов является фильтрация СП. Это связано с тем, что реальные спутниковые снимки обладают ярко выраженной пространственной неоднородностью. Использование для таких изображений стандартных оптимальных линейных фильтров приводит к значительным ошибкам. Решение данной проблемы возможно за счет применения специальных фильтров, учитывающих сложный характер изображений. Рассмотрим синтез таких фильтров для случая, когда можно не учитывать корреляционную связь между отдельными компонентами многокомпонентного изображения. В таком случае обработку компонент изображения можно проводить независимо друг от друга. Проведенные исследования [11-13] показывают, что для формирования таких фильтров можно использовать дважды стохастические модели изображений, позволяющие описывать неоднородные сигналы [14]. Рассмотрим в качестве примера следующую модель [8]:

$$x_{ijk} = 2\rho_{xij}x_{i-1,j,k} + 2\rho_{yij}x_{i,j-1,k} - 4\rho_{xij}\rho_{yij}x_{i-1,j-1,k} - \rho_{xij}^2x_{i-2,j,k} - \rho_{yij}^2x_{i,j-2,k} + 2\rho_{xij}^2\rho_{yij}x_{i-2,j-1,k} + 2\rho_{yij}^2\rho_{xij}x_{i-1,j-2,k} - \rho_{xij}^2\rho_{yij}^2x_{i-2,j-2,k} + b_{ij}\xi_{ijk} \quad (1)$$

где x_{ijk} - моделируемое СП с нормальным распределением $M\{x_{ijk}\}=0$, $M\{x_{ijk}^2\}=\sigma_x^2$; ξ_{ijk} - СП независимых стандартных гауссовских СВ $M\{\xi_{ijk}\}=0$, $M\{\xi_{ijk}^2\}=\sigma_\xi^2=1$; ρ_{xij} и ρ_{yij} - коэффициенты корреляции модели с кратными корнями характеристических уравнений кратности (2,2) [3]; b_{ij} - коэффициент масштаба моделируемого СП.

Случайные величины ρ_{xij} и ρ_{yij} с гауссовской плотностью распределения вероятности могут быть описаны следующими авторегрессионными уравнениями:

$$\begin{aligned} \tilde{\rho}_{xij} &= r_{1x}\tilde{\rho}_{x(i-1)j} + r_{2x}\tilde{\rho}_{xi(j-1)} - r_{1x}r_{2x}\tilde{\rho}_{x(i-1)(j-1)} + \sigma_{\rho_x}\sqrt{(1-r_{1x}^2)(1-r_{2x}^2)}\zeta_{\rho_xij}, \\ \tilde{\rho}_{yij} &= r_{1y}\tilde{\rho}_{y(i-1)j} + r_{2y}\tilde{\rho}_{yi(j-1)} - r_{1y}r_{2y}\tilde{\rho}_{y(i-1)(j-1)} + \sigma_{\rho_y}\sqrt{(1-r_{1y}^2)(1-r_{2y}^2)}\zeta_{\rho_yij}, \\ \rho_{xij} &= \tilde{\rho}_{xij} + m_{\rho_x}, \\ \rho_{yij} &= \tilde{\rho}_{yij} + m_{\rho_y}, \end{aligned} \quad (2)$$

где $r_{1x} = M\{\tilde{\rho}_{xij}\tilde{\rho}_{x(i-1)j}\}$, $r_{2x} = M\{\tilde{\rho}_{xij}\tilde{\rho}_{xi(j-1)}\}$ - коэффициенты корреляции случайного параметра $\tilde{\rho}_{xij}$; $r_{1y} = M\{\tilde{\rho}_{yij}\tilde{\rho}_{y(i-1)j}\}$, $r_{2y} = M\{\tilde{\rho}_{yij}\tilde{\rho}_{yi(j-1)}\}$ - коэффициенты корреляции случайного параметра $\tilde{\rho}_{yij}$; ζ_{ρ_xij} и ζ_{ρ_yij} - нормально распределенные СВ с $M\{\zeta_{\rho_xij}\} = M\{\zeta_{\rho_yij}\} = 0$, $M\{\zeta_{\rho_xij}^2\} = M\{\zeta_{\rho_yij}^2\} = \sigma_\zeta^2 = 1$.

Отметим, что модель (1) с параметрами (2) имитирует неоднородные изображения [14], что позволяет рекомендовать ее для описания реальных спутниковых изображений. При этом для фильтрации построеной развертки изображения, полученного в результате реализации такой модели, можно воспользоваться векторным (построчным) нелинейным фильтром Калмана [11,12]. Для этого объединим элементы строки изображения в вектор $\bar{x}_i = (x_{i1}, x_{i2}, \dots, x_{iN})^T$. Тогда модель отдельной компоненты изображения может быть записана в виде:

$$\bar{x}_i = \text{diag}(\bar{\rho}_{xi})\bar{x}_{i-1} + \nu(\bar{\rho}_{xi}, \bar{\rho}_{yi})\bar{\xi}_i, \quad \bar{\rho}_{xi} = r_{1x}\bar{\rho}_{x(i-1)} + \nu_{\rho_x}\bar{\xi}_{\rho_x}, \quad \bar{\rho}_{yi} = r_{1y}\bar{\rho}_{y(i-1)} + \nu_{\rho_y}\bar{\xi}_{\rho_y},$$

где $\text{diag}(\bar{\rho}_{xi})$ - диагональная матрица с элементами $\bar{\rho}_{xi}$ на главной диагонали; нижнетреугольная матрица ν - матрица, которая определяется разложением матрицы ковариаций: $V_x = \nu\nu^T$.

Процесс построчного оценивания описывается нелинейным фильтром Калмана:

$$\hat{x}_{pi} = \hat{x}_{spi} + P_i \frac{\partial \Phi^T}{\partial \bar{x}_{pi}} V_n^{-1} (\bar{z}_i - \hat{x}_{spi}), \quad \bar{x}_{pi} = \begin{pmatrix} \bar{x}_i \\ \bar{\rho}_{xi} \\ \bar{\rho}_{yi} \end{pmatrix} = \Phi(\bar{\rho}_{x(i-1)}\bar{x}_{i-1}) + \nu(\bar{\rho}_{x(i-1)}, \bar{\rho}_{y(i-1)})\bar{\xi}_i,$$

где $\bar{x}_{spi} = \Phi(\bar{x}_{p(i-1)})$, $\Phi_p(\bar{x}_{p(i-1)}) = \begin{pmatrix} \Phi(\rho, x) \\ r_{1x}\bar{\rho}_{x(i-1)} \\ r_{1y}\bar{\rho}_{y(i-1)} \end{pmatrix}$, $\bar{\xi}_i = \begin{pmatrix} \bar{\xi}_i \\ \bar{\xi}_{\rho_x} \\ \bar{\xi}_{\rho_y} \end{pmatrix}$, P_i - ковариационная матрица ошибок фильтрации.

Использование данного алгоритма возможно при условии точно известных характеристик информационного СП, т.е. коэффициентов r_{1x} , r_{2x} , r_{1y} , r_{2y} , а также ρ_{0x} , ρ_{0y} и $\sigma_{\rho_x}^2$, $\sigma_{\rho_y}^2$, σ_x^2 . В противном случае необходима предварительная оценка этих параметров. Для этого могут быть использованы псевдоградиентные процедуры [13,15] оценивания, а также выражения для КФ дважды стохастических моделей СП [14].

4. Результаты исследования эффективности обнаружения сигналов на реальных изображениях

Теперь сравним работу двух обнаружителей аномалий, построенных на базе дважды стохастической модели (Алгоритм 1) и на базе обычной авторегрессионной модели [2] (Алгоритм 2). При этом обнаружение будем выполнять на реальных изображениях, полученных со спутника LandSat-8. Исследования проводятся для трех изображений. При этом на каждом изображении выбрано по 4 области, где может находиться аномалия. Стоит отметить, что области выбираются исходя из структуры исследуемых изображений с учетом большей и меньшей неоднородности, а процедуры обнаружения выполняются не для всего изображения, а только для данных областей. На рис. 1а-1в показаны примеры изображений с сигналами, расположенными на разных участках изображений, а также отражены вероятности правильного обнаружения, полученные с использованием двух алгоритмов. Размеры всех изображений равны 250x250. При этом изображения искажаются белым гауссовым шумом с единичной дисперсией. Размер квадрата - 4x4, радиус круга - 2. Отношение сигнал/шум равно 1. Статистика снимается 150 раз.

В таблице 1 представлены величины выигрыша алгоритма 1 по отношению к алгоритму 2 по величине порогового сигнала при вероятности правильного обнаружения 0.5 и вероятности ложной тревоги 0.001, что соответствует порогу $L_0 = 3,1\sigma_{\varepsilon_1}^2$.

Таблица 1. Выигрыш (в процентах) алгоритма обнаружения на базе дважды стохастической модели по сравнению с алгоритмом обнаружения, построенном на базе АР модели

Фигура/Изображение	Позиция 1	Позиция 2	Позиция 3	Позиция 4
Квадрат на 1 изображении	0	0	0	0
Круг на 1 изображении	5	2	0	2
Квадрат на 2 изображении	68	3	13	4
Круг на 2 изображении	60	4	3	5
Квадрат на 3 изображении	21	4	4	5
Круг на 3 изображении	70	5	7	7

Анализ результатов, показывает, что алгоритм на базе дважды стохастической модели работает лучше алгоритма на базе простой авторегрессионной модели и обеспечивает уверенное обнаружение сигнала в 90-95% случаев. Малые значения выигрышей в таблице 1 объясняются тем фактом, что сигналы имеют малые размеры, и их окрестности в сопоставимом масштабе близки к однородным. Если же сигнал находится «на стыке» однородных областей, алгоритм на базе дважды стохастической модели обеспечивает значительный (до 70%) выигрыш по уровню сигнала. Выигрыши, представленные в таблице 1, рассчитаны для каждого случая из выражения

$$Gain = \frac{Pd_{ds} - Pd_{ar}}{Pd_{ar}},$$

где Pd_{ds} и Pd_{ar} - процент правильно обнаруженных сигналов на базе дважды стохастической и авторегрессионной моделей соответственно.

Из анализа рис. 1 можно сделать вывод о том, что аналогичная ситуация в смысле обеспечения как минимум одинаковой эффективности, а в ряде случаев и значительного выигрыша, сохраняется и при единичном отношении сигнал шум. При этом вероятность правильного обнаружения зависит не только от формы и размеров самого сигнала, но и от значений яркости в его ближайшей окрестности. В этом смысле, более универсальным алгоритмом выступает алгоритм на основе дважды стохастических моделей СП.

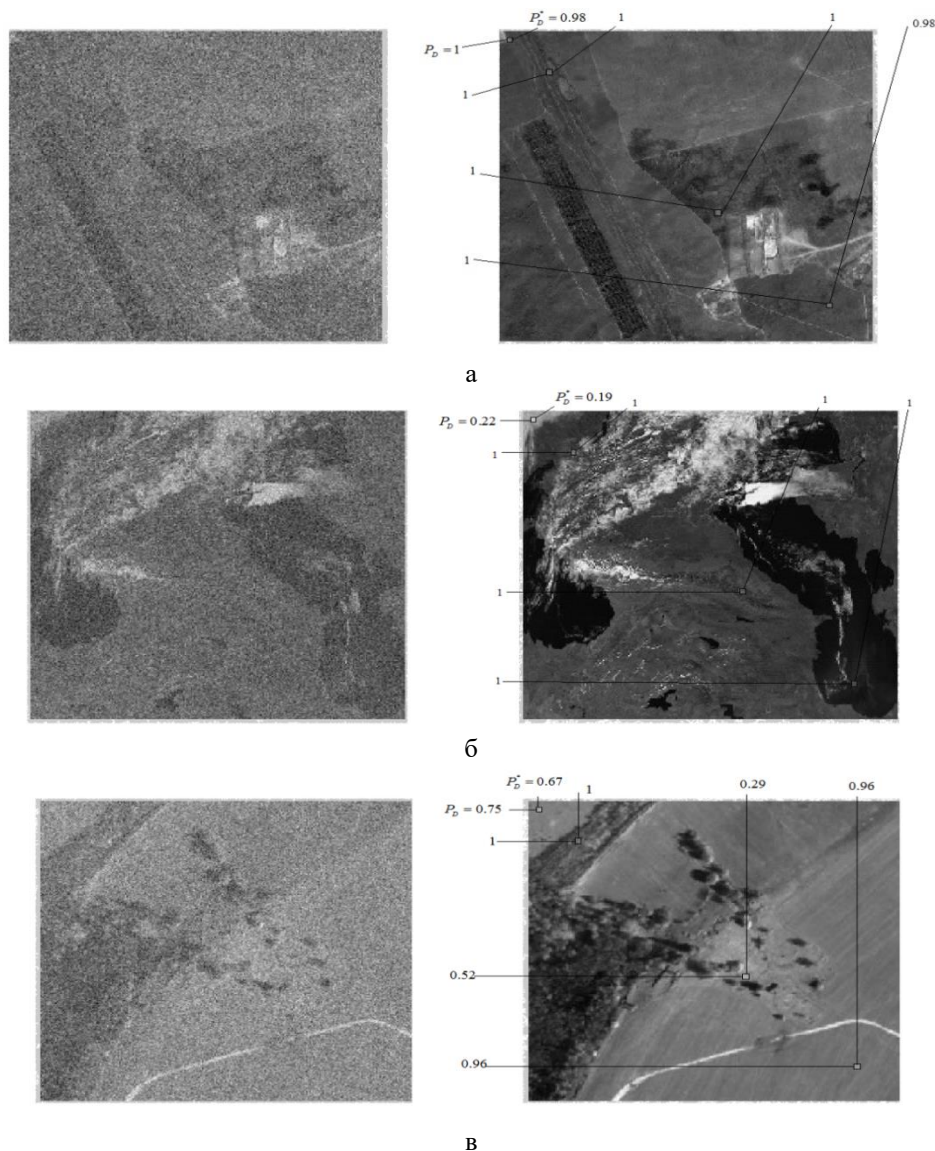


Рис.1. Зашумленные изображение (слева) и исходные изображения (справа) с вероятностями правильного обнаружения квадратного сигнала: слева приведены вероятности для Алгоритма 1, сверху – вероятности для Алгоритма 2.

5. Заключение

Проведен синтез и исследована эффективность правильного обнаружения на основе алгоритмов с применением дважды стохастических моделей СП. Статистическое моделирование показало, что алгоритм с использованием векторной калмановской фильтрации позволяет достигать значительных выигрышей по сравнению с алгоритмом на базе скалярной фильтрации в условиях имитации изображений на основе дважды стохастической модели СП. Основное преимущество векторной фильтрации заключается в возможности оценивания изменения параметров изображений. Разработанный алгоритм также применим к обнаружению протяженных сигналов на изображениях.

Благодарности

Работа выполнена при поддержке гранта РФФИ 16-41-732-027 "Построение стохастических моделей и алгоритмов обработки последовательностей неоднородных многозональных изображений для региональных систем экологического мониторинга".

Литература

- [1] Радиотехнические системы: учебник для студентов вузов / под ред. Ю.М. Казаринова. – М.: Издательский центр "Академия", 2008. – 592 с.
- [2] Перов, А.И. Статистическая теория радиотехнических систем: учебное пособие для вузов. / А.И. Перов. – М.: Радиотехника, 2003. – 400 с.
- [3] Васильев, В.К. Статистический анализ изображений. / К.К. Васильев, В.П. Крашенинников. – Ульяновск: УлГТУ, 2014. – 214 с.
- [4] Borghys, D. Hyperspectral anomaly detection: A comparative evaluation of methods / D. Borchys, V. Achard, S. R. Rotman, N. Gorelik, C. Perneel, E. Schweicher // General Assembly and Scientific Symposium, 2011, XXXth URSI. – 2011. – P. 1-4

- [5] Baghbidi, M. Z. Improvement of Anomaly Detection Algorithms in Hyperspectral Images Using Discrete Wavelet Transform / M.Z. Baghbidi, K. Jamshidi, A. R. Nilchi and S. Homayouni // *Signal & Image Processing: An International Journal (SIPIJ)*, Vol. 2, №4, December 2011, P. 13-25
- [6] Soofbaf, S. R. Efficient detection of anomalies in hyperspectral images / S.R. Soofbaf, M. J. Valadan Zoej, H. Fahimnejad, H. Ashoori // *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XXXVII. Part B7. Beijing 2008, P. 303-308.
- [7] Денисова, А.Ю. Обнаружение аномалий на гиперспектральных изображениях /А.Ю. Денисова, В.В. Мясников // *Компьютерная оптика*, т. 38, №2, 2014. - С. 287-296.
- [8] Vasil'ev, K.K. Doubly stochastic models of images / K.K. Vasil'ev, V.E. Dement'ev, N.A. Andriyanov // *Pattern Recognition and Image Analysis*, January 2015, vol. 25(1), P. 105-110. DOI: 10.1134/S1054661815010204
- [9] Васильев, К.К. Статистический анализ последовательностей многомерных изображений / К.К. Васильев, А.Г. Ташлинский, В.Р. Крашенинников // *Научные технологии*, 2013, №5, С. 5-11
- [10] Васильев, К.К. Оценивание параметров дважды стохастических случайных полей / К.К. Васильев, В.Е. Дементьев, Н.А. Андриянов // *Радиотехника*, 2014, №7, С. 103-106.
- [11] Васильев, К.К. Анализ эффективности оценивания изменяющихся параметров дважды стохастической модели / К. К. Васильев, В. Е. Дементьев, Н. А. Андриянов // *Радиотехника*. 2015. № 6. С. 12-15.
- [12] Васильев, К.К. Обнаружение протяженных сигналов на фоне дважды стохастических изображений / К. К. Васильев, В. Е. Дементьев, Н. А. Андриянов // *Радиотехника*, 2016, №9. - С. 23-27.
- [13] Vasil'ev, K.K. Application of mixed models for solving the problem on restoring and estimating image parameters / K.K. Vasil'ev, V.E. Dement'ev, N.A. Andriyanov // *Pattern Recognition and Image Analysis*. January 2016. V. 26(1). - P. 240-247. DOI: 10.1134/S1054661816010284
- [14] Андриянов, Н.А. Метод подгонки изображений на основе моделей случайных полей с изменяющимися параметрами // *Успехи современной науки*. 2016. Т. 5. №9. - С. 98 – 100
- [15] Андриянов, Н. А. Псевдоградиентные процедуры в задачах оценивания параметров моделей изображений // 26-я Международная Крымская конференция «СВЧ-техника и телекоммуникационные технологии» (КрыМиКо'2016). Севастополь, 4 – 10 сентября 2016 г.: материалы конф. в 2 т. – Севастополь, 2016. – Т. 1, С. 2705 – 2710