

УДК 519.21:528.85

doi:10.15217/issn1684-8853.2017.2.2

АНАЛИЗ ИНФОРМАТИВНЫХ ПРИЗНАКОВ В ЗАДАЧАХ ОБРАБОТКИ ДАННЫХ АЭРОКОСМИЧЕСКОГО МОНИТОРИНГА

В. И. Хименко^а, доктор техн. наук, профессор

М. Ю. Охтилев^{а, б}, доктор техн. наук, профессор

А. А. Ключарев^б, канд. техн. наук, доцент

В. А. Матяш^б, канд. техн. наук, доцент

^аОАО «НИО ЦИТ «Петрокомета», Санкт-Петербург, РФ

^бСанкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, РФ

Цель: анализ обобщенной модели аэрокосмического мониторинга, систематизация результатов по выбору информативных признаков в задачах обнаружения и распознавания исследуемых объектов, вероятностный анализ данных при многоспектральной структуре космических изображений. **Результаты:** на основе обобщенной структурной модели технологий космического мониторинга выделены характерные задачи, связанные с получением, преобразованием, обработкой и анализом информации при дистанционном зондировании Земли. Выполнена общая классификация наиболее распространенных информативных признаков применительно к задачам сегментации, обнаружения и распознавания объектов на космических изображениях. Показаны возможности вероятностного анализа спектральных индексов и особенности анализа многоспектральных изображений методами двумерных диаграмм рассеяния в пространстве интенсивностей. **Практическая значимость:** полученные результаты позволяют систематизировать и обобщить многочисленные разрозненные исследования по выбору и анализу информативных признаков. От выбора и структуры этих признаков существенно зависят процедуры синтеза алгоритмов обработки, сложность и общая эффективность решения основных задач космического мониторинга.

Ключевые слова — космический мониторинг, обработка информации, информативные признаки, распознавание изображений, вероятностный анализ, диаграммы рассеяния.

Введение

В настоящее время общий уровень развития аэрокосмических информационных технологий характеризуется широким спектром решаемых задач, наличием разного рода алгоритмов обработки данных и большим разнообразием самостоятельных областей практического (и потенциального) использования систем космического мониторинга. Технологии космического мониторинга позволяют получать обобщенную пространственно-временную информацию относительно объектов, явлений, процессов и состояния околоземной среды на наблюдаемых обширных территориях. Один и тот же космический снимок может служить источником информации для различных задач и областей научной и хозяйственной деятельности [1–5]: контроля окружающей среды и водных ресурсов, океанологии, геофизики и поиска полезных ископаемых, сельского хозяйства, лесного хозяйства и общего землепользования, мониторинга социально-экономического развития регионов, объектов техносферы, мониторинга стихийных бедствий, чрезвычайных ситуаций и т. д.

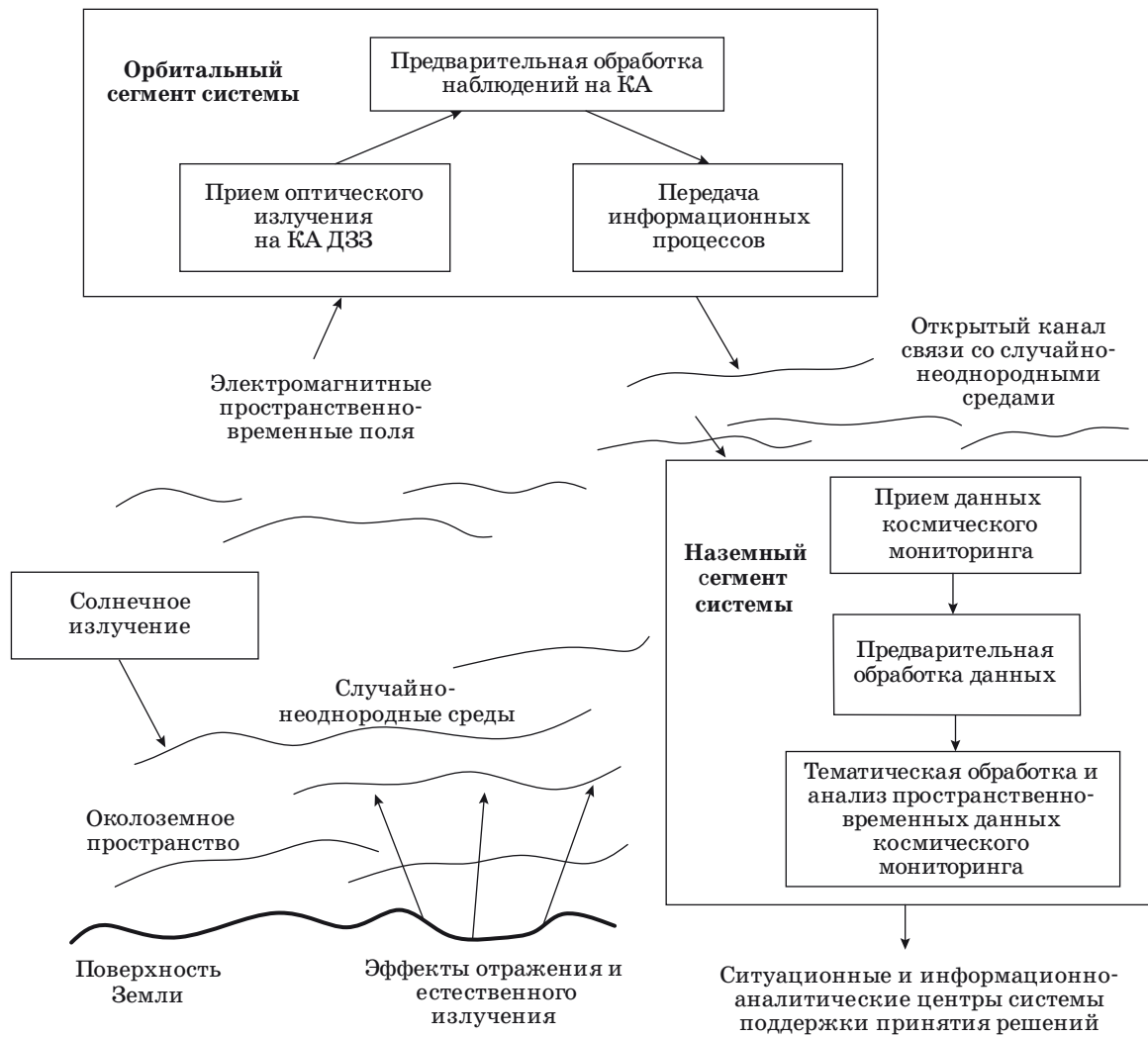
Подобное многообразие задач и междисциплинарность исследований всегда приводят к необходимости систематизации и обобщения накоплен-

ных результатов, разработки унифицированных структурных моделей систем и общей классификации наиболее важных информативных признаков, лежащих в основе синтеза алгоритмов обработки и анализа аэрокосмической информации. Рассмотрение перечисленных вопросов является основной целью данной работы.

Обобщенная модель получения и преобразования информации

Независимо от конкретной области исследований, основными задачами аэрокосмического мониторинга обычно являются задачи сбора, преобразования, передачи, обработки и анализа информации. Сбор информации выполняется космическими аппаратами (КА) дистанционного зондирования Земли (ДЗЗ), а передача данных на наземные станции приема осуществляется по специальным высокоскоростным каналам космической радиосвязи. Все основные операции мониторинга упрощенно можно представить в виде обобщенной структурной схемы (рис. 1).

Методы ДЗЗ относятся к классу косвенных методов исследования. При их реализации исследуемые параметры земной поверхности или околоземного пространства непосредственно не измеряются, и задачи мониторинга сводятся к ис-



■ Рис. 1. Обобщенная модель системы космического мониторинга

следованиям объектов, физических процессов и явлений по косвенным данным — косвенным измерениям.

Основным источником информации в системах дистанционного зондирования является электромагнитное излучение. Процессы формирования электромагнитных полей происходят за счет отраженного солнечного излучения (эффекты отражения) и собственного теплового излучения (эффекты естественного излучения) земной поверхности, различных объектов и сред, температура которых выше абсолютного нуля (см. рис. 1).

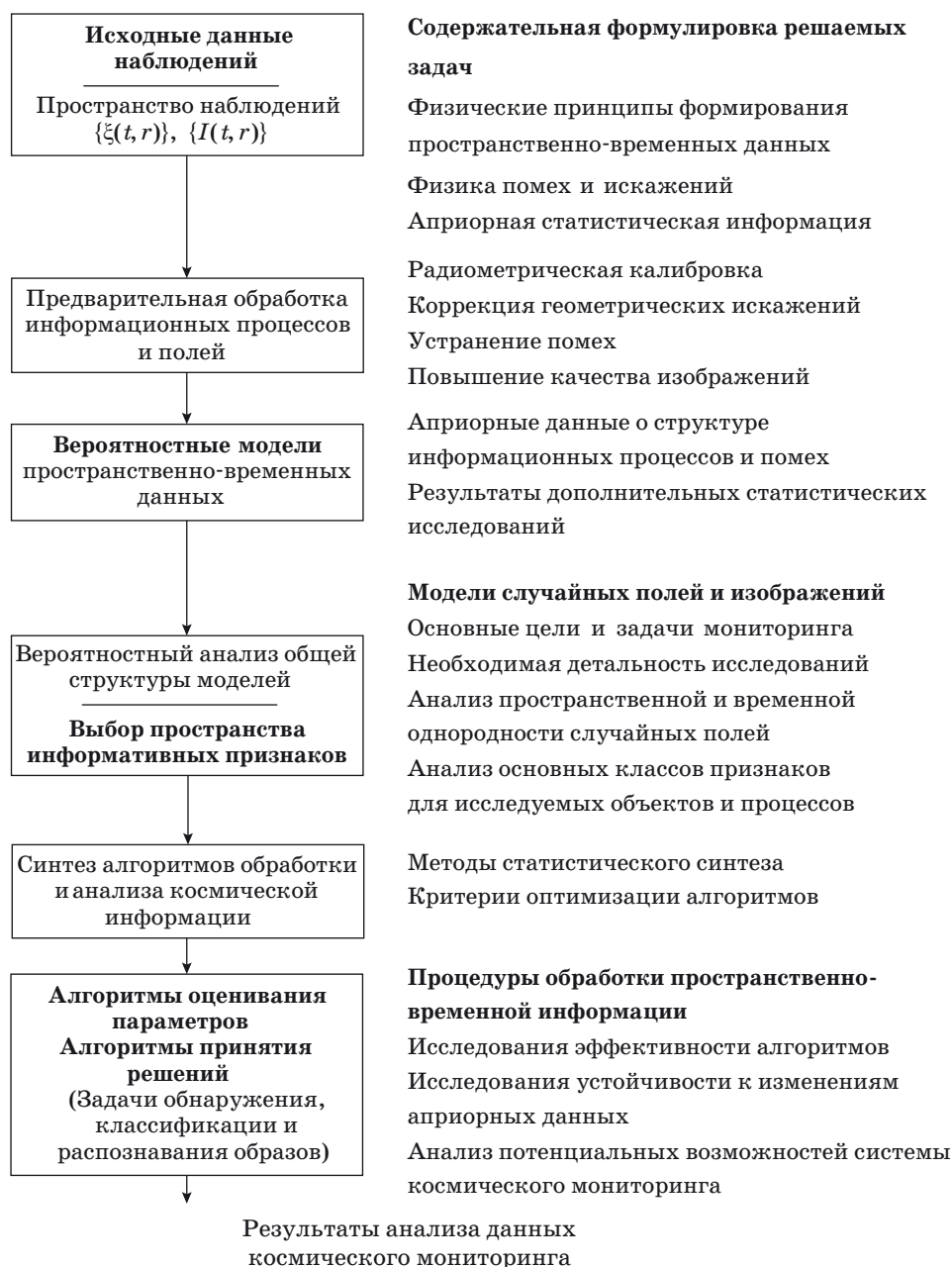
На КА выполняется прием электромагнитных излучений в заданном диапазоне спектра, проводится предварительная обработка и передача полученной информации на наземные станции приема. Все последующие операции связаны с тематической обработкой и анализом пространственно-временных данных на уровне наземных комплексов космического мониторинга.

При обработке полученных данных могут рассматриваться параметры интенсивности, спектральный состав, поляризация, направление распространения излучений. Эти характеристики функционально зависят от свойств, состояния и пространственного положения исследуемых объектов и среды распространения излучений. Использование подобных зависимостей лежит в основе всех методов аэрокосмического мониторинга и дает возможность по косвенным данным получить информацию о состоянии исследуемых объектов.

Структура тематической обработки и анализа данных

Рассмотрим теперь последовательность основных этапов тематической обработки данных и выделим характерные особенности задач, связанных с каждым из этих этапов. Общую струк-

Характерные особенности основных этапов обработки и анализа данных космического мониторинга



■ Рис. 2. Общая структура решения тематических задач мониторинга

туру обработки представим в виде обобщенной модели (рис. 2).

В системах космического мониторинга исходные данные на приемной стороне обычно имеют вид пространственно-временного случайного поля $\{\xi(t, r)\}$ и, как правило, при регистрации рассматриваются в форме двумерных изображений $I(x, y, t)$, где I — значение интенсивности (или яркости); x, y — координаты пространства $r = (x, y, z)$; t — время.

На этапе предварительной обработки основные операции связаны с радиометрической калибровкой данных, коррекцией геометрических искажений, устранением некоторых видов помех и повышением общего качества изображений. Многие из этих операций в настоящее время могут выполняться в автоматизированном виде [2, 6].

Все процедуры последующей обработки существенно зависят от содержания основных задач мониторинга и выбранных методов математиче-

ской формализации информационных процессов. При наиболее общем подходе описание пространственно-временных данных должно здесь выполняться на уровне вероятностных моделей случайных полей $\{\xi(t, r)\}$ и изображений $\{I(t, r)\}$. Важной особенностью подобного описания является то, что основные характеристики таких полей $\{\xi(t, r)\}$ изменяются во времени и в пространстве, и, следовательно, вероятностная структура случайных функций $\{\xi(t, r)\}$, $\{I(t, r)\}$ не обладает свойством пространственно-временной однородности.

После выбора и обоснования вероятностных моделей (см. рис. 2) необходимым этапом является процедура детального анализа их общих свойств и выбор информативных признаков для решения основных задач мониторинга. Этот этап является одним из наиболее важных и сложных. Выбор информативных признаков во многом определяет структуру алгоритмов обработки информации и общую эффективность решения задач, связанных с оцениванием параметров, обнаружением, классификацией и распознаванием исследуемых объектов на полученных изображениях.

Пространство информативных признаков

Задачи выбора информативных признаков возникают не только при обработке изображений в системах космического мониторинга. Они характерны для технической и медицинской диагностики, для систем технического зрения, систем интеллектуального управления и самых разнообразных систем распознавания образов. Вместе с тем проблема выбора информативных признаков в большинстве задач физики, техники, биологии и медицины общих решений не имеет [7–10].

С позиций общей теории распознавания образов основными целями перехода от полного описания исходных изображений к анализу пространства информативных признаков, в частности, являются:

- выделение характерных особенностей исследуемых объектов;
- максимизация различий между рассматриваемыми классами объектов;
- минимизация различий между объектами одного класса;
- упрощение описания рассматриваемых объектов и снижение размерности пространства представления данных;
- подготовка необходимых данных для решения задач синтеза алгоритмов оценивания параметров, алгоритмов обнаружения, классификации или идентификации отдельных объектов на изображениях.

Процедуры выбора признаков трудно поддаются формализации, они существенно зависят от содержания решаемых задач, априорных дан-

ных относительно рассматриваемых объектов и необходимой детальности исследований. К настоящему времени накоплено достаточно много эвристических результатов по описанию информативных признаков в задачах распознавания изображений. При этом появляется необходимость (и возможность) обобщения и рассмотрения общей классификации «различительных» признаков применительно к технологиям аэрокосмического мониторинга.

Все разнообразие существующих в этой области признаков условно можно разделить на несколько самостоятельных классов (рис. 3). Прежде всего, здесь можно выделить классы естественных и искусственных признаков [2, 5, 6, 11].

К классу естественных признаков обычно относятся характеристики, которые непосредственно отображаются на исследуемом изображении. В этот класс входят прямые признаки (яркостные, геометрические, структурные), косвенные признаки (индикаторное описание объектов, отдельных свойств, движений и изменений состояния) и различные виды комплексных признаков (совместное использование нескольких видов информации). При рассмотрении косвенных признаков обычно используются взаимосвязи между отдельными свойствами различных объектов, и информация об одном объекте получается из прямых признаков, относящихся к другому объекту.

Класс искусственных признаков включает в себя более сложные характеристики, нахождение которых связано со специальной обработкой или предварительным анализом изображений. К этому классу относится группа гистограммных признаков, спектральные и корреляционные признаки, аналитические и комплексные информативные признаки (см. рис. 3).

Для задач космического мониторинга наиболее распространенными и привычными являются яркостные признаки. Измерения таких признаков могут проводиться либо в отдельных точках, либо в окрестностях некоторых заданных областей. Можно при этом использовать сечения и проекции изображений, оценивать среднюю яркость отдельных фрагментов, можно перейти к классу искусственных признаков на основе некоторых выбранных линейных и нелинейных преобразований [6, 11–13].

Важную роль в задачах обнаружения и распознавания объектов играют геометрические признаки. В этот класс входят характеристики формы, размеры, площади, контурные особенности объектов. Контурные признаки обычно определяются резкими изменениями (перепадами, разрывами) яркости. На практике такие перепады яркости часто определяют очертания объектов, отображаемых на космических изображениях, и эта особенность заметно повышает эффективность кон-



■ **Рис. 3.** Классификация наиболее распространенных информативных признаков

турных признаков. Еще одна особенность связана с разнообразием возможных подходов к описанию и исследованию — при рассмотрении геометрических признаков могут использоваться метрические, топологические и аналитические методы.

В решении задач мониторинга основной интерес может быть связан с пространственной или временной структурой изображений, статическими или динамическими, детерминированными или вероятностными характеристиками. Выбранные при этом информативные признаки могут вычисляться для исследуемого изображения в целом или для отдельных фрагментов изображений, т. е. могут иметь глобальный или локальный характер. Все эти особенности должны отражаться в общей классификации пространства информативных признаков (см. рис. 3).

Особенности многоспектрального мониторинга

Современные системы космического мониторинга позволяют получать космические изображения одновременно в нескольких спектральных

диапазонах, т. е. относятся к классу многоспектральных или, при большом количестве каналов, к классу гиперспектральных [1, 2]. Увеличение числа спектральных каналов и расширение общего частотного диапазона исследований существенно повышает информативность систем мониторинга. Более того, при обработке многоспектральных данных появляется дополнительная возможность выбора наиболее информативных каналов — выбора тех спектральных диапазонов, в которых исследуемые объекты, процессы и явления наиболее полно проявляют свои характерные свойства.

Первичная информация на космических изображениях всегда отображается в значениях интенсивности (или яркости) $I(x, y | t, \lambda_i)$ каждого элемента — пикселя исследуемого изображения в определенный момент времени t в выбранном спектральном диапазоне $\lambda_i, i = 1, k$. При раздельном рассмотрении спектральных каналов обычно определяются типовые статистические характеристики: среднее значение яркости или математическое ожидание $m_I(\lambda_i)$, дисперсия $\sigma_I^2(\lambda_i)$, гистограмма распределения яркости $p^*(I; \lambda_i)$. Все подобные характеристики обладают простым и

наглядным физическим смыслом. Так, в частности, среднее значение яркости $m_I(\lambda_i)$ в каждом спектральном канале λ_i характеризует основные эффекты отражения и естественного излучения объектов в данном диапазоне спектра. Величина дисперсии $\sigma_I^2(\lambda_i)$ позволяет оценить диапазон изменений яркости и уровень контрастности изображения. Наиболее полной характеристикой является при этом гистограмма распределения яркости, которая дает возможность оценить общее распределение значений яркости по всему изображению, а также позволяет характеризовать степень однородности вероятностной структуры исследуемых пространственных данных.

При реализации алгоритмов совместной обработки многоспектральных данных наиболее часто для анализа выбираются два частотных диапазона, например λ_m и λ_n , $m \neq n$, рассматривается каждый пиксель изображения в одном диапазоне λ_m , оценивается его яркость и сравнивается с яркостью такого же пикселя в другом частотном диапазоне λ_n . На основе полученных данных можно определить простое отношение интенсивностей

$$S_{mn}(x, y) = I(x, y | \lambda_n) / I(x, y | \lambda_m), \quad (1)$$

которое позволяет перейти от рассмотрения абсолютных значений яркости к анализу относительных (или нормированных) значений. Подобные операции (1) лежат в основе синтеза новых преобразованных изображений и во многих задачах используются для формирования нового пространства спектральных признаков — индексного описания изображений [2–5]. К такому описанию, в частности, относится достаточно широкий класс вегетационных индексов VI (Vegetation Index) [2, 5]: RVI — относительный вегетационный индекс, PVI — ортогональный, NDVI — нормализованный разностный вегетационный индекс, SAVI — вегетационный индекс с коррективкой на почву, SARVI — вегетационный индекс с коррективкой на влияние почвы и атмосферы, TVI — группа трансформированных вегетационных индексов.

Существующее многообразие индексных описаний связано с разнообразием решаемых задач и попытками поиска наилучших спектральных признаков для каждой задачи. Вместе с тем нужно еще раз подчеркнуть, что информационной основой всех подобных характеристик являются простые отношения интенсивностей (1) для выбранных частотных диапазонов.

Вероятностный анализ спектральных признаков

Рассмотрим теперь возможности вероятностного анализа спектральных признаков в задачах обработки многоспектральных изображений.

В каждом спектральном диапазоне λ_i , $i = 1, 2, \dots, k$ значения интенсивностей $I(x, y | \lambda_i)$ изменяются в плоскости изображения (x, y) случайным образом. При рассмотрении двух произвольно выбранных диапазонов λ_m и λ_n , для полного вероятностного описания случайных величин $I(x, y | \lambda_m) = I_m$ и $I(x, y | \lambda_n) = I_n$ необходимо знать плотности вероятностей $p(I | \lambda_m) = p_m(I_m)$, $p(I | \lambda_n) = p_n(I_n)$ и совместную плотность вероятностей $p_{mn}(I_m, I_n)$. Если в подобной ситуации рассмотреть операцию перехода к индексному описанию изображений (1), то на основе общих правил функциональных преобразований [14, 15] для случайных значений индексов

$$\frac{I(x, y | \lambda_n)}{I(x, y | \lambda_m)} = \frac{I_n(x, y)}{I_m(x, y)} = \eta_{mn}(x, y) \quad (2)$$

может быть найдено общее выражение плотности вероятностей

$$p_\eta(\eta; x, y) = \int_{-\infty}^{\infty} p_{mn}(I_m, \eta I_m) |I_m| dI_m, \quad \eta = \eta_{mn}(x, y). \quad (3)$$

Если в рассматриваемых спектральных диапазонах λ_m , λ_n значения интенсивностей $I_m = I_m(x, y)$, $I_n = I_n(x, y)$ являются независимыми и характеризуются распределениями $p_m(I_m)$ и $p_n(I_n)$, то при совместном описании случайных величин I_m и I_n будет выполняться условие

$$p_{mn}(I_m, I_n) = p_m(I_m) p_n(I_n). \quad (4)$$

Общее выражение (3) примет при этом более простой вид

$$p_\eta(\eta; x, y) = \int_{-\infty}^{\infty} |I_m| p_m(I_m) p_n(\eta I_m) dI_m. \quad (5)$$

Следовательно, для определения вероятностных свойств спектральных отношений в подобных задачах достаточно знать одномерные распределения (4).

Приведенные результаты (2)–(5) позволяют выполнять общее исследование вероятностных характеристик для достаточно широкого класса спектральных индексов, определяемых на основе соотношений (1), (2). При этом важно подчеркнуть, что выражения (3) и (5) справедливы при различных распределениях $p_m(I_m; x, y)$, $p_n(I_n; x, y)$ значений интенсивностей $I_m = I(x, y | \lambda_m)$ и $I_n = I(x, y | \lambda_n)$. Вероятностная структура случайных величин $I_m(x, y)$ и $I_n(x, y)$ определяется здесь особенностями объектов космического мониторинга, содержанием решаемых задач и выбранными для анализа спектральными диапазонами λ_m, λ_n .

Как правило, на космических изображениях могут содержаться объекты нескольких различных классов. Каждый такой класс характеризуется относительно однородной вероятностной структурой данных. При обработке многозональных изображений в подавляющем большинстве задач обнаружения, распознавания и идентификации объектов для описания значений яркости в пределах каждого класса объектов в качестве типовой модели используется модель гауссова распределения [2–5].

Рассмотрим для этой наиболее распространенной модели особенности вероятностного описания спектральных индексов. Будем считать, что в спектральных каналах λ_m и λ_n значения интенсивностей $I_m(x, y)$ и $I_n(x, y)$ изменяются в пространстве (x, y) случайным образом и характеризуются совместно гауссовым распределением с параметрами:

— $m_I(\lambda_m) = M\{I_m(x, y)\}$, $m_I(\lambda_n) = M\{I_n(x, y)\}$ — математические ожидания (средние значения) интенсивностей в диапазонах λ_m и λ_n ;

— $\sigma_m^2 = \sigma_I^2(\lambda_m)$, $\sigma_n^2 = \sigma_I^2(\lambda_n)$ — дисперсии;

— $r_{mn} = R_{mn}(x, y) / \sigma_I(\lambda_m) \sigma_I(\lambda_n)$ — нормированное значение коэффициента корреляции $R_{mn}(x, y)$ для случайных величин $I_m(x, y)$ и $I_n(x, y)$.

Если в данном случае для упрощения преобразований рассмотреть центрированные значения интенсивностей

$$I_m^0 = I_m(x, y) - m_I(\lambda_m), \quad I_n^0 = I_n(x, y) - m_I(\lambda_n),$$

то на основе общей формулы (3) для отношения интенсивностей $\eta_{mn}^0(x, y) = I_n^0 / I_m^0$ получим

$$\begin{aligned} P_\eta(\eta^0; x, y) &= \int_{-\infty}^{\infty} P_{mn}(I_m^0, \eta^0 I_m^0) \Big| I_m^0 \Big| dI_m^0 = \\ &= \frac{1}{\pi \sigma_m \sigma_n (1 - r_{mn}^2)} \int_0^{\infty} I_m^0 \exp \left[-\frac{(I_m^0)^2}{2(1 - r_{mn}^2)} \times \right. \\ &\quad \left. \times \frac{\sigma_n^2 - 2r_{mn} \sigma_m \sigma_n \eta^0 + \sigma_m^2 (\eta^0)^2}{\sigma_m^2 \sigma_n^2} \right] dI_m^0 = \\ &= \frac{\sqrt{1 - r_{mn}^2}}{\pi \left[\frac{\sigma_n}{\sigma_m} - 2r_{mn} \eta^0 + \frac{\sigma_m}{\sigma_n} (\eta^0)^2 \right]}, \quad \eta^0 = \eta_{mn}^0(x, y). \quad (6) \end{aligned}$$

Этот результат позволяет исследовать вероятностную структуру спектральных индексов (2) в задачах двумерной обработки данных при гауссовых распределениях отдельных компонент. Достаточно наглядно проявляется здесь и влияние возможной корреляционной зависимости между значениями интенсивностей $I_m(x, y)$ и $I_n(x, y)$.

В частности, если в выбранных для анализа спектральных диапазонах λ_m и λ_n полученные данные мониторинга являются некоррелированными, то параметр $r_{mn} = 0$ и общее выражение (6) приводят к более простому результату

$$P_\eta(\eta^0; x, y) = \frac{\sigma_m \sigma_n}{\pi \left[\sigma_n^2 + \sigma_m^2 (\eta^0)^2 \right]}, \quad \eta^0 = \eta_{mn}^0(x, y). \quad (7)$$

Важно подчеркнуть, что полученное распределение (7) по своему виду совпадает с известным распределением Коши [16]. Такой результат характерен для исследования операций деления при независимых гауссовых случайных величинах.

На практике в задачах анализа спектральных индексов (2) плотность вероятностей (7) удобнее представить для значений $\eta = \eta_{mn}(x, y)$ в стандартной форме распределения Коши:

$$\begin{aligned} P_\eta(\eta; x, y) &= \frac{b}{\pi \left[b^2 + (\eta - a)^2 \right]}, \\ a &= \frac{m_I(\lambda_n)}{m_I(\lambda_m)}, \quad b = \frac{\sigma_n}{\sigma_m}. \quad (8) \end{aligned}$$

Функция (8) обладает свойством унимодальности и является симметричной относительно оси $\eta = a$. Значение a соответствует здесь параметру расположения и характеризует моду и медиану распределения (8). Параметр b является параметром масштаба.

Следует отметить, что для классического распределения Коши при изменениях переменной в диапазоне $(-\infty, \infty)$ моментные функции не существуют [16]. Это относится даже к параметрам математического ожидания и дисперсии. В задачах обработки изображений в качестве переменной рассматривается интенсивность (или яркость), случайные изменения которой ограничены заданным диапазоном $I(x, y) \in [I_{\min}, I_{\max}]$. Эта особенность приводит к использованию вероятностных моделей (3), (5)–(8) в условиях усеченных значений переменной $\eta_{mn}(x, y)$ и, следовательно, в условиях обеспечения конечности моментных функций. В частности, для усеченных распределений (3), (5)–(8) при любом выборе спектральных диапазонов λ_m, λ_n реальные значения

$$m_\eta(x, y) = M\{\eta_{mn}(x, y)\} < \infty, \quad \sigma_\eta^2 = \sigma_\eta^2(x, y) < \infty.$$

Анализ данных на основе диаграмм рассеяния

При обработке многоспектральных изображений всегда возникают вопросы, связанные с анализом вероятностной структуры и оценкой вза-

имной коррелированности данных, вопросы отображения и наглядного представления информации. В решении подобных вопросов принципиально могут использоваться различные методы многомерного анализа [16, 17], однако в качестве одного из наиболее простых и наглядных подходов целесообразно выделить метод, основанный на рассмотрении диаграмм рассеяния.

Предположим, что для обработки данных космического мониторинга используются изображения, полученные в двух спектральных диапазонах λ_m и λ_n . При раздельном рассмотрении этих изображений для каждого спектрального канала могут быть построены гистограммы распределения яркости. Пример такой гистограммы $p^*(I_m; x, y)$ для диапазона λ_m показан на рис. 4. По своей форме она близка к модели гауссова распределения и дает достаточно полное описание характера вероятностного распределения значений яркости I_m изображения $I_m(x, y | \lambda_m)$ в пространстве (x, y) .

Если перейти к совместному анализу структуры изображений в двух спектральных диапазонах λ_m и λ_n , то можно воспользоваться представлением двумерной информации в виде диаграммы рассеяния [2, 4, 18]. Для этого на плоскости значений яркости (I_m, I_n) формируется диаграмма, каждая точка которой отражает яркость отдельного элемента (пикселя) изображения $I_m(x, y | \lambda_m)$ в диапазоне λ_m (ось абсцисс — OI_m) и значение яркости такого же элемента изображения $I_n(x, y | \lambda_n)$ в диапазоне λ_n (ось ординат — OI_n). Полученная таким образом диаграмма характеризует особенности общего распределения значений яркости изображений $I_m(x, y)$ и $I_n(x, y)$ в двух выбранных для анализа спектральных каналах λ_m и λ_n .

В качестве иллюстрации на рис. 5 показан характерный вид подобной диаграммы рассеяния для двух изображений $I_m(x, y)$ и $I_n(x, y)$ с некорре-

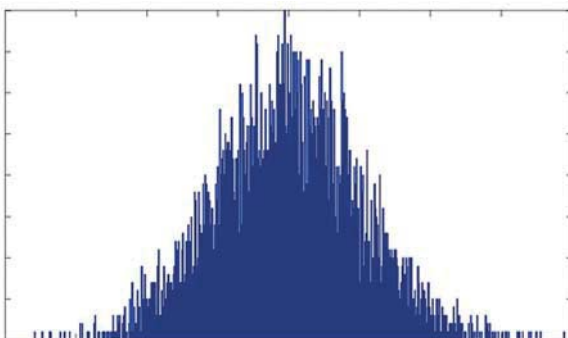
лированными и совместно гауссовыми распределениями значений яркости I_m, I_n в пространстве (x, y) . При формировании этой диаграммы в канале λ_m для наглядности использованы данные изображения $I_m(x, y)$, описываемого гистограммой $p^*(I_m; x, y)$, представленной на рис. 4.

Вероятностная структура изображения $I_n(x, y)$ в спектральном диапазоне λ_n аналогична структуре изображения $I_m(x, y)$, однако отличается меньшей величиной дисперсии $\sigma_n^2 = \sigma_I^2(\lambda_n)$ — меньшим разбросом значений яркости I_n по оси ординат OI_n . Все эти особенности наглядно проявляются на плоскости (I_m, I_n) в структуре двумерной диаграммы рассеяния (см. рис. 5).

В дополнение к этому типовому варианту на рис. 6 показаны еще две диаграммы рассеяния с характерной для анализа изображений вероятностной структурой. На одной из них (рис. 6, а) приведен пример проявления корреляционной зависимости между значениями яркости I_m и I_n изображений в рассматриваемых диапазонах λ_m и λ_n . Кроме того, из этой диаграммы видно, что в обоих спектральных диапазонах на изображениях присутствуют некоторые элементы, отличающиеся более высокой яркостью по сравнению со средним уровнем яркости основного изображения.

Другая диаграмма (рис. 6, б) демонстрирует характерный пример анализа изображений с неоднородной вероятностной структурой. Одна совокупность данных описывает здесь элементы изображений с взаимно коррелированными и относительно малыми уровнями яркостей I_m и I_n . Вторая группа характеризует наличие на изображениях объектов с высокой яркостью и практически некоррелированной структурой. Наглядно видна здесь и возможность разделения двух классов объектов на рассматриваемых изображениях по уровню их яркости.

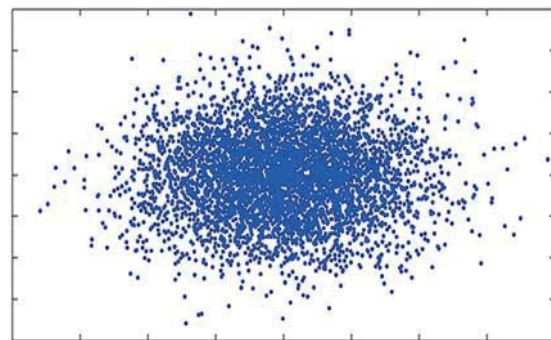
$p^*(I_m; x, y)$



Уровень яркости изображения I_m

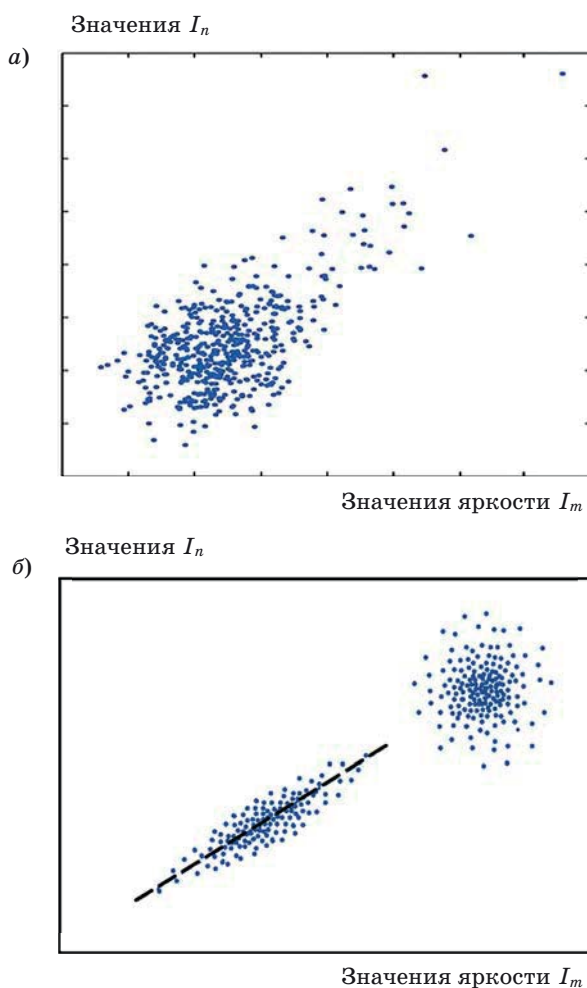
■ Рис. 4. Гистограмма распределения значений яркости изображения $I_m(x, y | \lambda_m)$ в пространстве (x, y)

Значения I_n



Значения яркости I_m в канале λ_m

■ Рис. 5. Типовая структура двумерной диаграммы рассеяния при некоррелированных гауссовых компонентах $I_m(x, y), I_n(x, y)$ и $\sigma_m^2 > \sigma_n^2$



■ **Рис. 6.** Характерный вид диаграмм рассеяния при появлении корреляционной зависимости (а) и при неоднородной вероятностной структуре изображений (б)

В целом необходимо отметить, что представление информации в виде двумерных диаграмм рассеяния позволяет в простой и наглядной форме отображать на плоскости (I_m, I_n) особенности вероятностной структуры изображений $I_m(x, y)$ и $I_n(x, y)$ и особенности их взаимосвязей.

Заключение

Представленные в данной работе результаты позволяют на уровне обобщенных моделей (см. рис. 1 и 2) выделить основные задачи вероятностного анализа и синтеза для систем аэрокосмического мониторинга на этапах получения, преобразования и обработки информационных процессов. В работе рассмотрены вопросы выбора информативных признаков при решении задач обнаружения и распознавания исследуемых объектов на космических изображениях, предло-

жен один из возможных подходов к общей классификации пространства признаков, выполнен вероятностный анализ спектральных индексных представлений информации и на основе двумерных диаграмм рассеяния показаны возможности совместной обработки данных при анализе многоспектральных изображений.

В целом представленные результаты позволяют сделать некоторые общие выводы.

1. Необходимость построения и анализа обобщенных моделей получения, преобразования и тематической обработки информации связана с возможностями систематизации накопленных результатов и возможностью разработки единого подхода к решению различных по своему содержанию задач космического мониторинга.

2. Задачи выбора информативных признаков характерны для любых систем обнаружения, классификации или распознавания объектов, процессов и явлений. Эти задачи трудно формализуются, в их решении существенную роль играют эвристические методы и использование накопленных статистических данных.

3. В зависимости от содержания решаемых задач один и тот же информативный признак исследуемого объекта может иметь высокую эффективность для одних задач и может быть малоэффективным в задачах другого класса. К сожалению, во многих реальных ситуациях набор информативных признаков определяется не тем, что нужно было бы измерить для решения конкретной задачи, а тем, что удастся измерить.

4. При обработке и анализе многоспектральных изображений достаточно часто выполняется совместная обработка данных различных спектральных каналов и на основе некоторых линейных и нелинейных преобразований вводятся новые информативные признаки. Полезно при этом подчеркнуть, что подобные преобразования не могут приводить к увеличению информативности исследуемых многоспектральных изображений. Такие операции могут улучшать визуальное восприятие данных, повышать наглядность отображения отдельных областей на изображениях при решении задач обнаружения или классификации заданных объектов.

5. Анализ пространственно-временной вероятностной структуры изображений на основе двумерных диаграмм рассеяния дает возможность накапливать, усреднять и систематизировать выборочные данные, выполнять визуальную оценку функциональных зависимостей, проверять однородность и выявлять неоднородные по яркости фрагменты в пространстве исследуемых изображений. Помимо этого, диаграммы рассеяния дают полезную информацию для решения типовых задач сегментации и распознавания на изображениях объектов различных классов.

Литература

1. Лебедев В. В., Гансвинд И. Н. Проектирование систем космического мониторинга. — М.: Наука, 2010. — 388 с.
2. Schowengerdt R. A. Remote Sensing: Models and Methods for Image Processing. — N. Y.: Elsevier, 2007. — 560 p.
3. Chandra A. M., Ghosh S. K. Remote Sensing and Geographical Information System. — New Delhi: Narosa Publishing House, 2006. — 312 p.
4. Rees W. G. Physical Principles of Remote Sensing. — Cambridge: Cambridge University Press, 2001. — 336 p.
5. Книжников Ю. Ф., Кравцова В. И., Тутубалина О. В. Аэрокосмические методы географических исследований. — М.: Академия, 2011. — 416 с.
6. Методы компьютерной обработки изображений / под ред. В. А. Соифера. — М.: Физматлит, 2003. — 784 с.
7. Васильев В. И. Распознающие системы: справочник. — Киев: Наукова думка, 1983. — 424 с.
8. Verhagen C., Duin R., Groen F. Progress Report on Pattern Recognition // Reports on Progress in Physics. 1980. Vol. 43. N 6. P. 785–831.
9. Потапов А. С. Распознавание образов и машинное восприятие. — СПб.: Политехника, 2007. — 548 с.
10. Top J. T., Gonzalez R. C. Pattern Recognition Principles. — London: Addison-Wesley Publishing Company, 1974. — 412 p.
11. Лабутина И. А. Дешифрование аэрокосмических снимков. — М.: Аспект Пресс, 2004. — 184 с.
12. Gonzales R. C., Woods R. E. Digital Image Processing. — New Jersey: Prentice Hall, 2002. — 1072 p.
13. Shapiro L. G., Stokman G. C. Computer Vision. — New Jersey: Prentice Hall, 2001. — 752 p.
14. Тихонов В. И. Статистическая радиотехника. — М.: Радио и связь, 1982. — 624 с.
15. Тихонов В. И., Харисов В. Н. Статистический анализ и синтез радиотехнических устройств и систем. — М.: Радио и связь, 2004. — 608 с.
16. Вероятность и математическая статистика: энциклопедия / под ред. Ю. В. Прохорова. — М.: БРЭ, 2003. — 910 с.
17. Ивченко Г. И., Медведев Ю. И. Введение в математическую статистику. — М.: ЛКИ, 2010. — 600 с.
18. Хименко В. И. Диаграммы рассеяния в анализе случайных потоков событий // Информационно-управляющие системы. 2016. № 4. С. 85–93. doi:10.15217/issn1684-8853.2016.4.85

UDC 519.21:528.85

doi:10.15217/issn1684-8853.2017.2.2

Informative Feature Analysis in Data Processing for Aerospace Monitoring

Khimenko V. I.^a, Dr. Sc., Tech., Professor, vih.13@yandex.ruOkhtilev M. Yu.^{a,b}, Dr. Sc., Tech., Professor, oxt@mail.ruKlucharev A. A.^b, PhD., Tech., Associate Professor, ak@aanet.ruMatiash V. A.^b, PhD., Tech., Associate Professor, vam@aanet.ru^aJSC «The Scientific Research and Experimental Centre of Intelligent Technologies «Petrokometa», 67, B. Morskaja St., 190000, Saint-Petersburg, Russian Federation^bSaint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaja St., 190000, Saint-Petersburg, Russian Federation

Purpose: We analyze a generalized model of aerospace monitoring, systematize the informative feature selection results in the problems of object detection and identification, and perform probabilistic data analysis with multispectral structure of space images. **Results:** On the base of a generalized structural model of space monitoring technologies, some specific problems have been specified in relation to obtaining, transforming, processing and analyzing the information from Earth remote sensing. The most common informative features have been generally classified in relation to the problems of satellite image object segmentation, detection and recognition. We have shown the possibilities of probabilistic analysis of spectral indices and the features of multispectral image analysis by two-dimensional scattering patterns in a space of intensities. **Practical relevance:** The obtained results allow you to systematize and generalize numerous disparate studies of the selection and analysis of informative features. The processing algorithm synthesis procedures, the complexity and general efficiency of solving the basic space monitoring problems significantly depend on the choice and structure of these features.

Keywords — Space Monitoring, Data Processing, Informative Features, Image Recognition, Probabilistic Analysis, Scatterplot.

References

1. Lebedev V. V., Gansvind I. N. *Proektirovanie sistem kosmicheskogo monitoringa* [Remote Sensing Systems Design]. Moscow, Nauka Publ., 2010. 388 p. (In Russian).
2. Schowengerdt R. A. *Remote Sensing: Models and Methods for Image Processing*. New York, Elsevier, 2007. 560 p.
3. Chandra A. M., Ghosh S. K. *Remote Sensing and Geographical Information System*. New Delhi, Narosa Publishing House, 2006. 312 p.
4. Rees W. G. *Physical Principles of Remote Sensing*. Cambridge, Cambridge University Press, 2001. 336 p.
5. Knizhnikov U. F., Kravtsova V. I., Tutubalina O. V. *Aerokosmicheskie metody geograficheskikh issledovanii* [Aerospace Methods of Geographical Research]. Moscow, Academia Publ., 2011. 416 p. (In Russian).
6. *Metody kompiuternoii obrabotki izobrazhenii* [Methods of Computer Image Processing]. Ed. V. A. Soifer. Moscow, Fizmatlit Publ., 2003. 784 p.
7. Vasiliev V. I. *Raspoznaiushchie sistemy* [Recognizing Systems. Handbook]. Kiev, Naukova dumka Publ., 1983. 424 p. (In Russian).

8. Verhagen C., Duin R., Groen F. Progress Report on Pattern Recognition. *Reports on Progress in Physics*, 1980, vol. 43, no 6, pp. 785–831.
9. Potapov A. S. *Raspoznavanie obrazov i mashinnoe vospriatie* [Pattern Recognition and Machine Perception]. Saint-Petersburg, Politekhnik Publ., 2007. 548 p. (In Russian).
10. Ton J. T., Gonzalez R. C. *Pattern Recognition Principles*. London, Addison-Wesley Publishing Company, 1974. 412 p.
11. Labutina I. A. *Deshifrovaniye aerokosmicheskikh snimkov* [Aerospace Images Decoding]. Moscow, Aspekt Press Publ., 2004. 184 p. (In Russian).
12. Gonzales R. C., Woods R. E. *Digital Image Processing*. New Jersey, Prentice Hall, 2002. 1072 p.
13. Shapiro L. G., Stokman G. C. *Computer Vision*. New Jersey, Prentice Hall, 2001. 752 p.
14. Tikhonov V. I. *Statisticheskaya radiotekhnika* [Statistical Radiotechnics]. Moscow, Radio i sviaz' Publ., 1982. 624 p. (In Russian).
15. Tikhonov V. I., Kharisov V. N. *Statisticheskii analiz i sintez radiotekhnicheskikh ustroystv i sistem* [Statistical Analysis and Synthesis of Radio Engineering Devices and Systems]. Moscow, Radio i sviaz' Publ., 2004. 608 p. (In Russian).
16. Prokhorov Yu. V. *Veroiatnost' i matematicheskaya statistika: entsiklopediya* [Probability and Mathematical Statistics: Encyclopedia]. Moscow, BRE Publ., 2003. 910 p. (In Russian).
17. Ivchenko G. I., Medvedev Yu. I. *Vvedenie v matematicheskuyu statistiku* [Introduction to Mathematical Statistics]. Moscow, LKI Publ., 2010. 600 p. (In Russian).
18. Khimenko V. I. Scatterplots to Analysis of Random Stream of Events. *Informatsionno-upravlyaiushchie sistemy* [Information and Control Systems], 2016, no. 4, pp. 85–93 (In Russian). doi:10.15217/issn 1684-8853.2016.4.85

Уважаемые авторы!

При подготовке рукописей статей необходимо руководствоваться следующими рекомендациями.

Статьи должны содержать изложение новых научных результатов. Название статьи должно быть кратким, но информативным. В названии недопустимо использование сокращений, кроме самых общепринятых (РАН, РФ, САПР и т. п.).

Объем статьи (текст, таблицы, иллюстрации и библиография) не должен превышать эквивалента в 20 страниц, напечатанных на бумаге формата А4 на одной стороне через 1,5 интервала Word шрифтом Times New Roman размером 13, поля не менее двух сантиметров.

Обязательными элементами оформления статьи являются: индекс УДК, заглавие, инициалы и фамилия автора (авторов), ученая степень, звание (при отсутствии — должность), полное название организации, аннотация и ключевые слова на русском и английском языках, электронные адреса авторов, которые по требованию ВАК должны быть опубликованы на страницах журнала. При написании аннотации не используйте аббревиатур и не делайте ссылок на источники в списке литературы.

Статьи авторов, не имеющих ученой степени, рекомендуются публиковать в соавторстве с научным руководителем, наличие подписи научного руководителя на рукописи обязательно; в случае самостоятельной публикации обязательно предоставляйте заверенную по месту работы рекомендацию научного руководителя с указанием его фамилии, имени, отчества, места работы, должности, ученого звания, ученой степени — эта информация будет опубликована в ссылке на первой странице.

Формулы набирайте в Word, не используя формульный редактор (MathType или Equation), при необходимости можно использовать формульный редактор; для набора одной формулы не используйте два редактора; при наборе формул в формульном редакторе знаки препинания, ограничивающие формулу, набирайте вместе с формулой; для установки размера шрифта никогда не пользуйтесь вкладкой Other..., используйте заводские установки редактора, не подгоняйте размер символов в формулах под размер шрифта в тексте статьи, не растягивайте и не сжимайте мышью формулы, вставленные в текст; в формулах не отделяйте пробелами знаки: + = -.

Для набора формул в Word никогда не используйте Конструктор (на верхней панели: «Работа с формулами» — «Конструктор»), так как этот ресурс предназначен только для внутреннего использования в Word и не поддерживается программами, предназначенными для изготовления оригинал-макета журнала.

При наборе символов в тексте помните, что символы, обозначаемые латинскими буквами, набираются светлым курсивом, русскими и греческими — светлым прямым, векторы и матрицы — прямым полужирным шрифтом.

Иллюстрации предоставляются отдельными исходными файлами, поддающимися редактированию:

— рисунки, графики, диаграммы, блок-схемы предоставляйте в виде отдельных исходных файлов, поддающихся редактированию, используя векторные программы: Visio 4, 5, 2002-2003 (*.vsd); Coreldraw (*.cdr); Excel (*.xls); Word (*.doc); AdobeIllustrator (*.ai); AutoCad (*.dxf); Matlab (*.ps, *.pdf или экспорт в формат *.ai);

— если редактор, в котором Вы изготавливаете рисунок, не позволяет сохранить в векторном формате, используйте функцию экспорта (только по отношению к исходному рисунку), например, в формат *.ai, *.esp, *.wmf, *.emf, *.svg;

— фото и растровые — в формате *.tif, *.png с максимальным разрешением (не менее 300 pixels/inch).

Наличие подрисовочных подписей обязательно (желательно не повторяющих дословно комментарии к рисункам в тексте статьи).

В редакцию предоставляются:

— сведения об авторе (фамилия, имя, отчество, место работы, должность, ученое звание, учебное заведение и год его окончания, ученая степень и год защиты диссертации, область научных интересов, количество научных публикаций, домашний и служебный адреса и телефоны, e-mail), фото авторов: анфас, в темной одежде на белом фоне, должны быть видны плечи и грудь, высокая степень четкости изображения без теней и отблесков на лице, фото можно представить в электронном виде в формате *.tif, *.png с максимальным разрешением — не менее 300 pixels/inch при минимальном размере фото 40×55 мм;

— экспертное заключение.

Список литературы составляется по порядку ссылок в тексте и оформляется следующим образом:

— для книг и сборников — фамилия и инициалы авторов, полное название книги (сборника), город, издательство, год, общее количество страниц;

— для журнальных статей — фамилия и инициалы авторов, полное название статьи, название журнала, год издания, номер журнала, номера страниц;

— ссылки на иностранную литературу следует давать на языке оригинала без сокращений;

— при использовании web-материалов указывайте адрес сайта и дату обращения.

Список литературы оформляйте двумя отдельными блоками на сайте журнала (<http://i-us.ru/paperrules>) по разным стандартам: Литература — СИБИД РФ, References — один из мировых стандартов.

Более подробно правила подготовки текста с образцами изложены на нашем сайте в разделе «Оформление статей».

Контакты

Куда: 190000, Санкт-Петербург,
Б. Морская ул., д. 67, ГУАП, РИЦ

Кому: Редакция журнала «Информационно-управляющие системы»

Тел.: (812) 494-70-02

Эл. почта: i.us.spb@gmail.com

Сайт: www.i-us.ru