

Классификация водных объектов по спектрам, измеренным с борта МКС в космическом эксперименте «Ураган»

М. Ю. Беляев¹, Б. И. Беляев², Л. В. Катковский², А. О. Мартинов²,
Э. Э. Сармин¹, О. О. Слюк², А. В. Чумаков²

¹ *Ракетно-космическая корпорация «Энергия» имени С. П. Королёва
г. Королёв, 141070, Россия
E-mail: post@rsce.ru*

² *Институт прикладных физических проблем имени А. Н. Севченко Белорусского
государственного университета, Минск, 220045, Республика Беларусь
E-mail: remsens@mail.ru*

В статье описаны результаты, полученные на основе обработки данных разновременных измерений изображений и спектров земной поверхности научной аппаратурой «Фотоспектральная система» (ФСС) с борта Международной космической станции по программе космического эксперимента «Ураган». Разработана новая методика классификации спектральных данных на основе метода главных компонент и корреляционного анализа, эффективная для классификации объектов, спектры отражения которых близки между собой как по форме, так и по абсолютным значениям, каковыми, в частности, являются зарегистрированные из космоса спектры водных акваторий различных морей и океанов ввиду очень малого вклада в них излучения водной поверхности и преобладающего вклада атмосферы. В отличие от традиционного применения преобразования главных компонент, когда для анализа оставляют несколько первых главных компонент, в предлагаемой методике исключается из анализа первая компонента (имеющая наибольшую дисперсию) и выполняется обратное преобразование в спектральное пространство, что позволяет учесть незначительные различия в спектрах объектов. Кроме того, вместо трудной для восприятия визуализации кластеров спектров в многомерном спектральном пространстве предложена наглядная визуализация классифицированных спектров в трёхмерном пространстве, построенном на основе взаимных корреляций спектров выборки. С использованием предложенной методики проведена классификация водных объектов земного шара (морей и океанов) по данным ФСС, которая продемонстрировала возможность дискриминации водных объектов по их спектрам отражения, что свидетельствует о существовании характерных спектральных особенностей акваторий различных морей.

Ключевые слова: дистанционное зондирование Земли, спектры, корреляционная методика, метод главных компонент, тематическая классификация

Одобрена к печати: 01.10.2019

DOI: 10.21046/2070-7401-2019-16-6-201-208

Введение

Фотоспектральная система (ФСС), разработанная в Институте прикладных физических проблем имени А. Н. Севченко Белорусского государственного университета по заказу ракетно-космической корпорации «Энергия» имени С. П. Королева и предназначенная для одновременной регистрации спектров и изображений земной поверхности с борта Международной космической станции (МКС), работает с 2010 г. на борту МКС по программе «Экспериментальная отработка наземно-космической системы мониторинга и прогноза развития природных и техногенных катастроф» космического эксперимента «Ураган» (Беляев и др., 2014а).

Система ФСС состоит из двух модулей: модуля регистрации изображения (МРИ) и модуля спектрорадиометра (МС). Измеренные спектрорадиометром спектры представляют собой абсолютные значения спектральной плотности энергетической яркости (СПЭЯ) в диапазоне длин волн от 350 до 1050 нм с шагом по спектру менее 1 нм и со спектральным разрешением не хуже 3–5 нм. Данные МРИ представляют собой изображения в формате RAW в трёх

цветовых каналах R, G, B без радиометрической коррекции (Беляев и др., 2014б). При съёмке с МКС один кадр цветного изображения охватывает площадку $\sim 35\,000 \times 25\,000$ м с пространственным разрешением около 10 м (проекция пикселя на поверхность Земли), тогда как проекция входной щели спектро радиометра на поверхность Земли имеет размер порядка 50×6000 м. Для каждого изображения спектро радиометр регистрирует три спектра с фиксированными временными интервалами между ними.

Анализ спектров отражения водных поверхностей. Корреляционная методика классификации спектров

Описываемая методика классификации по спектральным данным разрабатывалась специально для классификации объектов с близкими спектрами отражения во всём исследуемом спектральном диапазоне, каковыми, в частности, являются зарегистрированные ФСС с борта МКС спектры различных морей и океанов. С использованием географических координат орбиты станции во время съёмок были отобраны следующие объекты для анализа: спектры и изображения Каспийского, Красного, Средиземного морей и Атлантического океана. Всего для классификации было отобрано 239 спектров водных объектов. Все отобранные значения СПЭЯ были организованы в виде прямоугольной матрицы $\{x_{ij}\}$, в которой по столбцам располагаются различные спектры (признак или характеристика объекта, номер столбца соответствует номеру спектра), а по строкам — значения СПЭЯ в зависимости от длины волны от 350 до 950 нм (номер строки соответствует номеру спектрального канала).

В спектрах водных поверхностей, полученных из космоса, превалирует вклад рассеянного атмосферой излучения (атмосферная дымка) и лишь единицы процентов составляет вклад излучения, отражённого водной поверхностью и выходящего из неё, которое только и содержит информацию об объекте интереса. Полученные спектральные кривые имеют практически идентичную форму, что затрудняет использование известных методов тематической классификации. В свою очередь, на оба указанных вклада (атмосферы и воды) значительное влияние оказывают оптические параметры атмосферы и условия освещения, в частности зенитный угол Солнца. На рис. 1 показаны примеры измерений ФСС из космоса спектров водных поверхностей при безоблачной атмосфере в различных зонах планеты (моря, океаны, крупные внутренние водоёмы) при зенитных углах Солнца от 40 до 70° . Как видим, СПЭЯ водных поверхностей сильно зависит от зенитного угла Солнца, но длина волны максимума λ_{\max} восходящего излучения смещается незначительно и находится в диапазоне $410\text{--}415$ нм.

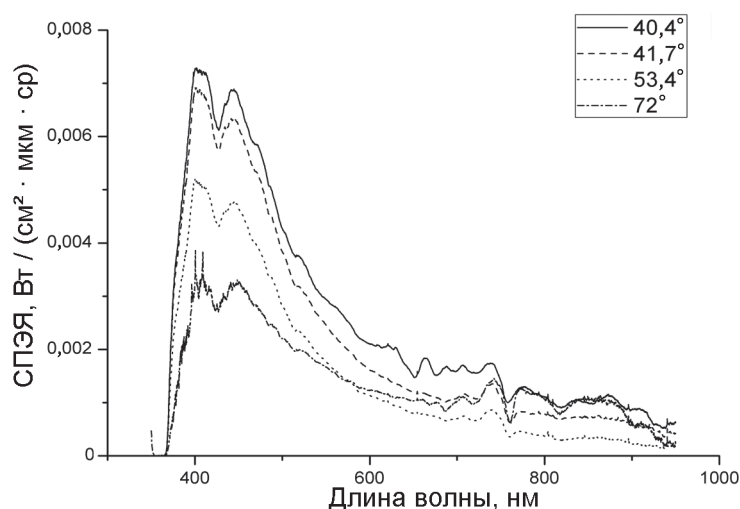


Рис. 1. Экспериментальные спектры восходящего излучения, измеренные ФСС из космоса над водными поверхностями при различных зенитных углах Солнца

Одним из наиболее известных способов устранения факторов, влияющих на спектр объекта, но не информативных для задачи классификации, является преобразование главных компонент. В настоящей работе мы его используем для устранения доминирующего фактора, преобладающего во всех спектрах водных поверхностей и делающего их похожими.

Предварительно, до указанного преобразования, для устранения разброса абсолютных значений спектров (для нас важна форма спектральной кривой) по каждому столбцу матрицы была выполнена нормировка исходных спектров следующего вида:

$$x'_{kn} = \frac{x_{kn} - \bar{x}_n}{\sigma(x_n)},$$

где \bar{x}_n и $\sigma(x_n)$ — математическое ожидание и среднеквадратическое отклонение по n -му столбцу, $k = 1, \dots, K$ и $n = 1, \dots, N$; K — количество спектральных каналов в спектре; N — общее количество спектров в выборке.

Далее над спектрами выполняется преобразование главных компонент, что целесообразно, если спектры имеют схожую форму. Метод главных компонент преобразовывает систему зависимых между собой значений СПЭЯ в спектральных каналах к системе статистически независимых главных компонент, по которым можно разложить каждый из исходных спектров. Считается, что каждая главная компонента является отображением независимых между собой свойств снимаемых объектов в исследуемой спектральной области.

Типичные спектры водных поверхностей представлены на *рис. 2*, а вид первой главной компоненты спектров водных поверхностей показан на *рис. 3*.

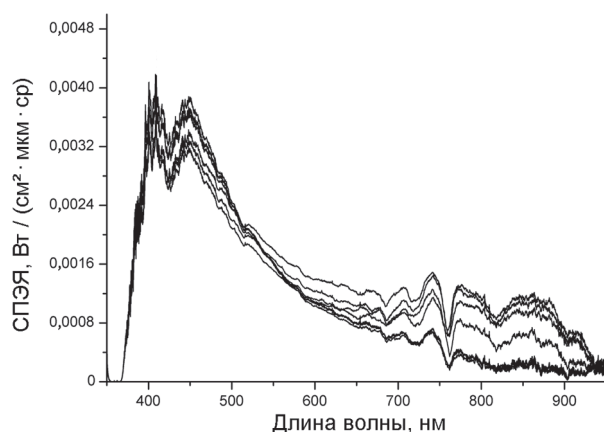


Рис. 2. Спектры отражения различных морей с борта МКС

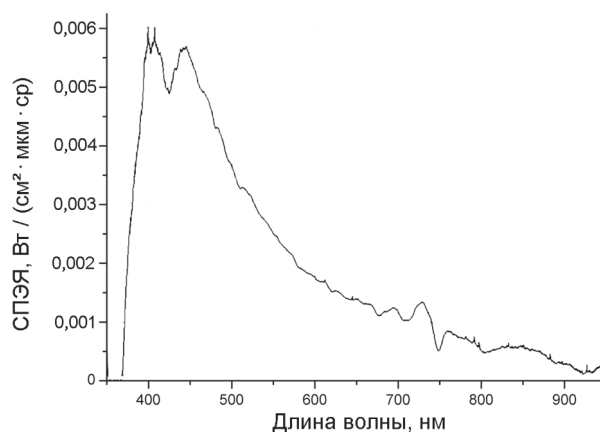


Рис. 3. Первая главная компонента исследуемых спектров водных поверхностей

Первая главная компонента вносит максимальный вклад в дисперсию выборки (соответствует наиболее интенсивному фактору в отражённом излучении исследуемого объекта). В случае со спектрами водных поверхностей, полученных из космоса, это, очевидно, вклады спектра освещающего излучения Солнца и ослабления излучения атмосферой.

Корректировка спектров с учётом первой главной компоненты может выполняться с использованием двух вариантов: нормировка на первую главную компоненту или её зануление.

Матрица главных компонент \mathbf{Z} (размерностью $K \times K$) строится на основе матрицы весовых коэффициентов $\mathbf{A} = \{a_{kn}\}$, $k = 1, \dots, K$ (число каналов), $n = 1, \dots, N$ (число спектров), учитывающих тесноту связи между исходными признаками (спектрами) \mathbf{X} и главными компонентами:

$$\mathbf{Z} = \mathbf{X} \times \mathbf{A}.$$

В нашем случае для устранения влияния главного фактора использовано обнуление столбца первой главной компоненты в матрице главных компонент, которое, как выяснилось, даёт лучший результат (в сравнении с нормировкой на неё). Над модифицированной

таким образом матрицей \mathbf{Z}' выполняется обратное преобразование в исходное спектральное пространство и восстанавливаются спектры:

$$\mathbf{X} = \mathbf{Z}' \times \mathbf{A}^{-1}. \quad (1)$$

Дальнейшая классификация (кластеризация) строится с использованием коэффициента корреляции r_{nm} полученных таким образом новых спектров, поскольку и у преобразованных спектров основную информацию несёт форма кривой:

$$r_{nm} = \frac{\sum_{k=1}^K (x_{kn} - \bar{x}_n) \cdot (x_{km} - \bar{x}_m)}{\sqrt{\sum_{k=1}^K (x_{kn} - \bar{x}_n)^2 \sum_{k=1}^K (x_{km} - \bar{x}_m)^2}}, \quad \bar{x}_n = \frac{1}{K} \sum_{k=1}^K x_{kn}, \quad \bar{x}_m = \frac{1}{K} \sum_{k=1}^K x_{km}, \quad (2)$$

где $\{x_{kn}\}$ и $\{x_{km}\}$ — два различных спектра n и m из всего набора; k — номер спектрального канала.

Однако визуализация разделения спектров на кластеры или классы (при любом методе классификации) в многомерном спектральном пространстве трудна для восприятия (обычно, как, например, в ENVI (Environment for Visualizing Images), это осуществляется путём отображения проекций облака точек в двумерное или трёхмерное пространство). Для визуализации разделения спектров на классы в случае, когда в качестве метрики используется коэффициент корреляции (2), вместо многомерного спектрального пространства предлагается следующий способ визуализации в трёхмерном пространстве, который выступает одновременно как способ классификации и визуализации. По коэффициентам взаимной корреляции определяются три наиболее различающихся по форме (коэффициенту корреляции) спектра. Вначале находятся два спектра с номерами s и t , которые соответствуют минимальному элементу корреляционной матрицы r_{st} . Третий спектр k определяется из выражения:

$$\min_{n=1, N} (r_{ns} + r_{nt}), \quad (3)$$

т.е., n соответствует номеру спектра, для которого сумма в выражении (3) минимальна для уже найденных s и t .

Найденные таким образом три спектра используются для формирования трёх осей прямоугольной системы координат. Для отображения в этой трёхмерной системе координат точек, соответствующих всем остальным спектрам выборки, вдоль каждой из осей откладываются значения коэффициента корреляции текущего спектра с соответствующим спектром (s , t , n). Матрица взаимных корреляций определяет, таким образом, координаты спектров в указанной трёхмерной системе координат. На расположение спектра влияет его степень сходства с тремя спектрами, выбранными для формирования осей. Пример диаграммы корреляционных коэффициентов спектров показан на *рис. 4* (см. с. 205), где по осям откладываются значения коэффициента корреляции спектра r с тремя наиболее различающимися спектрами, которые на каждом этапе алгоритма кластеризации выбираются автоматически и в данном случае являются 7-м, 38-м и 199-м столбцами в сформированной матрице $\{x_{ij}\}$ из 239 спектров. Объекты со схожими спектральными кривыми, а значит, и оптико-физическими свойствами, кластеризуются. На *рис. 4* и далее спектры, отнесённые к разным кластерам, отображаются на диаграммах точками разных цветов.

После применения корреляционного способа кластеризации без выполнения коррекции главных компонент диаграмма корреляционных коэффициентов спектров имеет вид, представленный на *рис. 5* (см. с. 205) (показана двумерная диаграмма). После процедуры коррекции главных компонент наименее коррелирующие между собой спектры являются 2-м и 245-м столбцом матрицы $\{x_{ij}\}$. На основе этой диаграммы был исключён одиночный спектр в правом нижнем углу (соответствующий по виду СПЭЯ облачности), после чего диаграмма приобрела вид, показанный на *рис. 6* (см. с. 205).

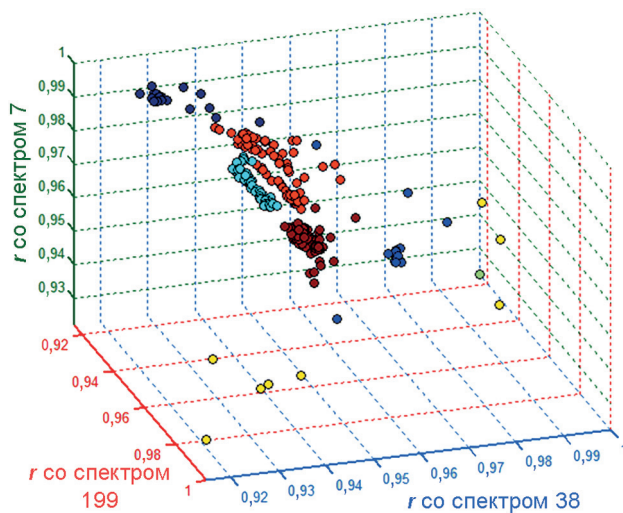


Рис. 4. Диаграмма корреляционных коэффициентов спектров водных поверхностей

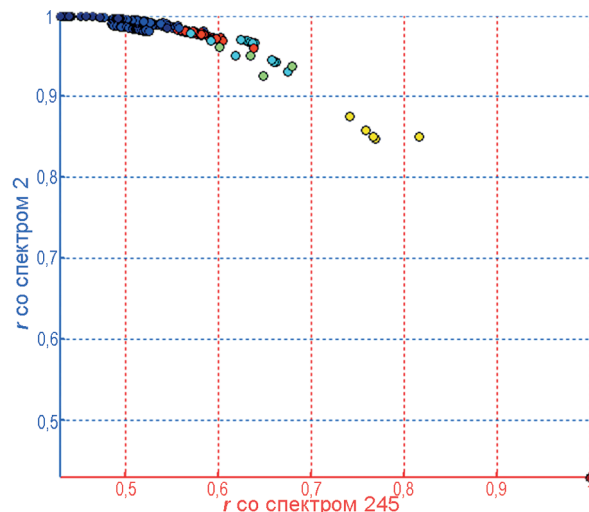


Рис. 5. Диаграмма корреляционных коэффициентов спектров водных поверхностей

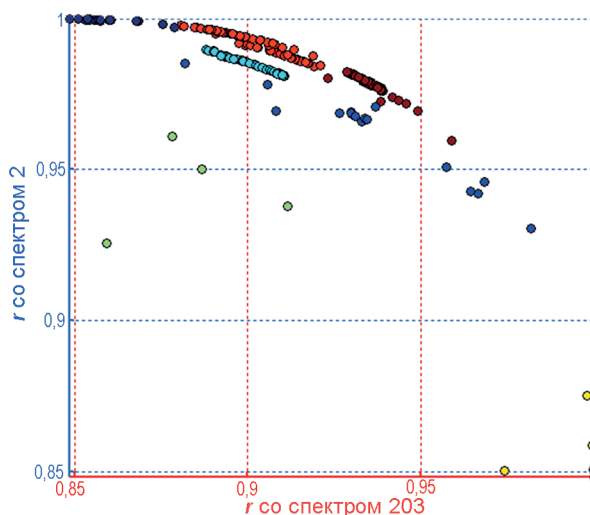


Рис. 6. Диаграмма корреляционных коэффициентов спектров водных поверхностей после удаления спектра облаков

Последняя диаграмма также позволила выявить и исключить из рассмотрения несколько нетипичных для воды спектров (в правом нижнем углу), которые, как установлено по сопутствующим изображениям, принадлежат смеси песка и воды (рис. 7, см. с. 206).

Для оставшихся спектров, имеющих почти идентичную форму (см. рис. 2), применено преобразование главных компонент с устранением первой главной компоненты.

Полученные после обратного преобразования (1) спектры имеют уже визуальную различимую форму, их типичные представители для разных акваторий показаны на рис. 8 (см. с. 206).

Для модифицированных спектров диаграмма по значениям взаимных корреляций с тремя выделенными спектрами (осями) имеет вид, показанный на рис. 9 (см. с. 207). Новая диаграмма спектров характеризуется возросшей дисперсией значений в матрице корреляций.

Далее было проведено разделение всей выборки (239 спектров) на 7 классов с использованием иерархического агломеративного метода (Лагутин, 2007). Преимуществом иерархических методов кластеризации является их наглядность. Выбор в качестве метрики связывания центров кластеров Уорда и меры сходства стандартизированного евклидова расстояния дало лучший результат кластеризации (рис. 10, см. с. 207).

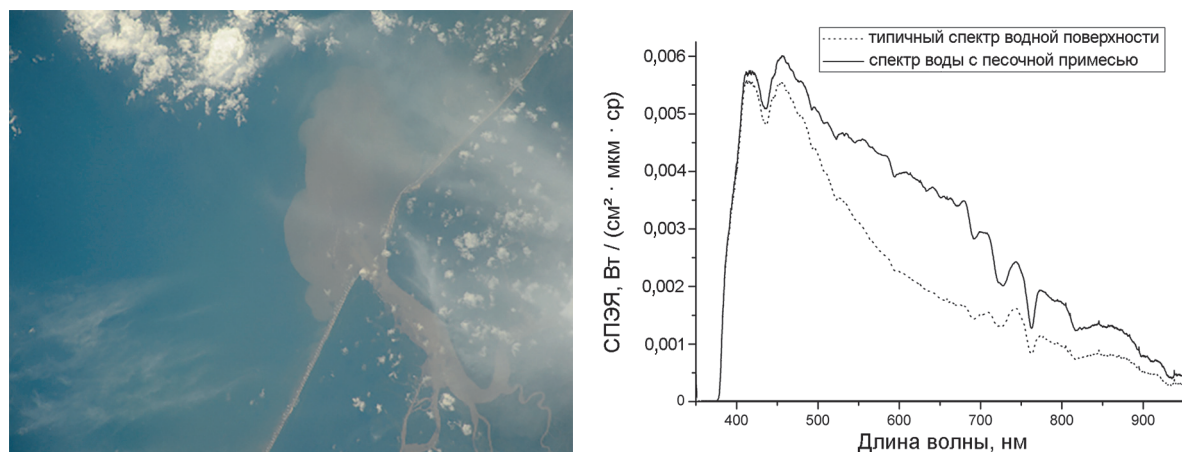


Рис. 7. Вынос воды с песочной примесью из устья р. Нигер в Атлантический океан (слева) и соответствующий спектр в сравнении с типичным спектром водной поверхности (справа), зарегистрированные ФСС

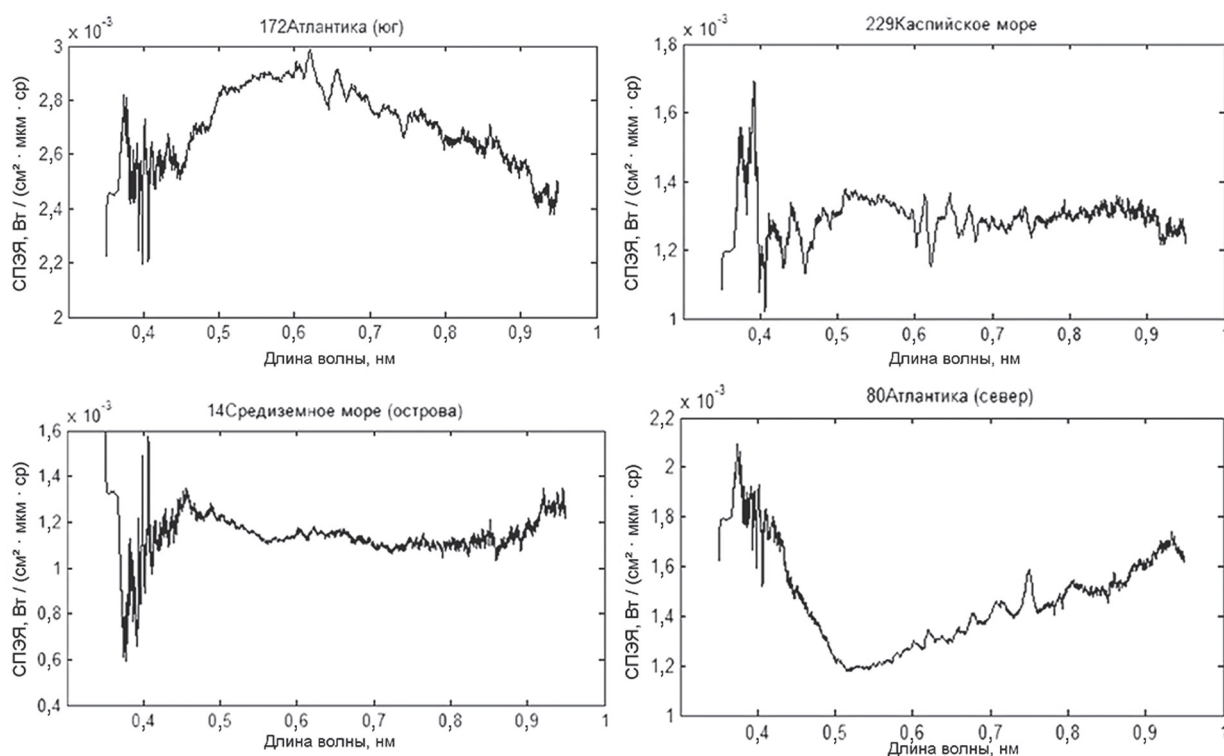


Рис. 8. Спектральные кривые, восстановленные после устранения влияния первой главной компоненты

Номера классов 1–7, которыми обозначены кластеры на рис. 10 (см. с. 207), соответствуют следующим водным объектам, установленным по координатно-временной привязке съёмки при движении МКС по орбите:

- 1) Атлантический океан с высокой облачностью;
- 2) Каспийское море;
- 3) Эгейское море;
- 4) Средиземное море (рядом с о. Крит);
- 5) Средиземное море;
- 6) Атлантический океан;
- 7) Красное море.

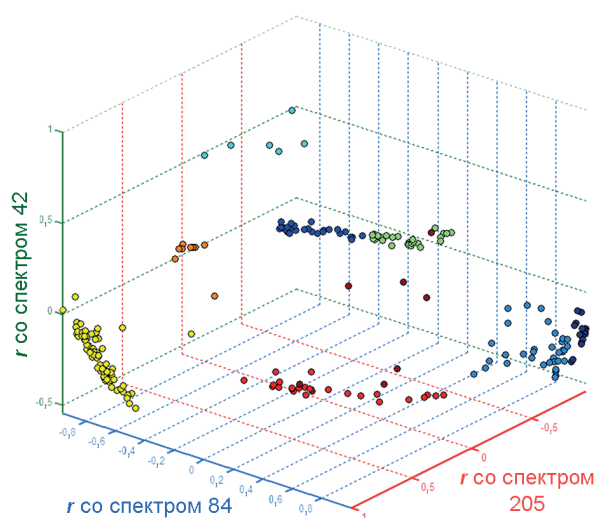


Рис. 9. Диаграмма корреляционных коэффициентов спектров без учёта первой главной компоненты

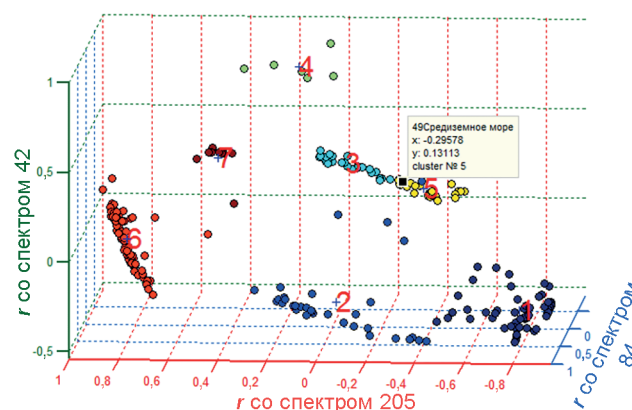


Рис. 10. Разделение выборки спектров на классы

По результатам кластеризации правильно было распознано 212 спектров из 239 (88,7 %).

Заключение

Разработана корреляционная методика тематической классификации объектов по их спектрам отражения, которая применена к спектрам водных акваторий, снятых ФСС с борта МКС, а также подтверждена эффективность данной методики. Её применение в совокупности с коррекцией данных с помощью преобразования главных компонент позволяет решить задачу классификации выборок спектров, очень близких по форме и абсолютным значениям, для которых не работают традиционные алгоритмы. Предложен также наглядный способ визуализации результатов классификации в трёхмерном пространстве (вместо многомерного спектрального пространства). Результаты исследования спектров водных объектов свидетельствуют о существовании характерных спектральных особенностей водных акваторий различных морей, позволяющих проводить их идентификацию.

К достоинствам предложенной методики можно отнести: простоту интерпретации результатов, так как можно выполнять даже визуальную классификацию; возможность повышения качества классификации на основе знания физических особенностей природных объектов; независимость от выбора критерия (метрики) сравнения спектров (суть методики не изменится при замене коэффициента корреляции на другую меру сравнения); возможность распространения на другие задачи, не связанные со спектрами. Методика может быть использована для данных любой аппаратуры дистанционного зондирования Земли с произвольным спектральным диапазоном и числом спектральных каналов.

Литература

1. Беляев Б. И., Беляев М. Ю., Денисов Л. В., Катковский Л. В., Крот Ю. А., Сармин Э. Э. (2014а) Результаты испытаний фотоспектральной системы на МКС // Исслед. Земли из космоса. 2014. № 6. С. 27–39.
2. Беляев Б. И., Беляев М. Ю., Десинев Л. В., Катковский Л. В., Сармин Э. Э. (2014б) Обработка спектров и изображений с фотоспектральной системы в космическом эксперименте «Ураган» на МКС // Исслед. Земли из космоса. 2014. № 6. С. 54–65.
3. Лагутин Б. М. Наглядная математическая статистика: учеб. пособие. М.: БИНОМ. Лаборатория знаний, 2007. 472 с.

Classification of water surfaces by spectra measured from the ISS during the Uragan space experiment

M. Yu. Belyaev¹, B. I. Belyaev², L. V. Katkovsky², A. O. Martinov²,
E. E. Sarmin¹, V. A. Siliuk², A. V. Chumakov²

¹ *S. P. Korolev Rocket and Space Public Corporation "Energia", Korolev 141070, Russia
E-mail: post@rsce.ru*

² *A. N. Sevchenko Institute of Applied Physical Problems of Belorussian State University
Minsk 220045, Republic of Belarus
E-mail: remsens@mail.ru*

The paper describes the results of processing of multi-temporal measurements (images and spectra) of the Earth surface made by the scientific equipment Photospectral System (PhSS) from the International Space Station (ISS) during the Uragan space experiment. A new classification methodology for spectral data based on the principal component method and correlation analysis has been developed. The methodology is effective for the classification of objects whose reflection spectra are close to each other both in form and in absolute values. In contrast to the traditional application of principal component transformation, where the first few principal components are used for analysis, in the proposed method the first component that has the greatest variance is excluded from the analysis. The inverse transformation into the spectral space is performed, which allows taking into account minor differences in the spectra of objects. In addition, a visualization of the classified spectra in three-dimensional space based on the correlations of the entire sample of spectra is proposed. Using the proposed methodology, classifications of the water surfaces (seas and oceans) are done on the basis of PhSS data. The results of classification demonstrate the possibility of dividing water surfaces according to their reflection spectra, which indicates the existence of characteristic spectral features of the waters of different seas.

Keywords: remote sensing, spectra, correlation method, principal component analysis, classification

Accepted: 01.10.2019

DOI: 10.21046/2070-7401-2019-16-6-201-208

References

1. Belyaev B. I., Belyaev M. Yu., Desinov L. V., Katkovskii L. V., Krot Yu. A., Sarmin E. E. (2014a), Rezul'taty ispytaniy fotospektral'noi sistemy na MKS (Photospectral system test results on ISS), *Issledovanie Zemli iz kosmosa*, 2014, No. 6, pp. 27–39.
2. Belyaev B. I., Belyaev M. Yu., Desinov L. V., Katkovskii L. V., Sarmin E. E. (2014b), Obrabotka spektrov i izobrazhenii s fotospektral'noi sistemy v kosmicheskom eksperimente "Uragan" na MKS (Spectra and images processing from Photospectral system in space experiment "Hurricane" on the ISS), *Issledovanie Zemli iz kosmosa*, 2014, No. 6, pp. 54–65.
3. Lagutin B. M., *Naglyadnaya matematicheskaya statistika* (Visible mathematical statistics), Moscow: BINOM. Laboratoriya znaniy, 2007, 472 p.