

УДК 629.197

В. Н. МАСЛЕЙ¹, К. Г. БЕЛОУСОВ¹, Д. К. МОЗГОВОЙ², В. С. ХОРОШИЛОВ¹,
А. С. БУШАНСКАЯ², Д. В. МАСЛЕЙ², Д. А. ХМАРА¹

¹ Государственное предприятие «Конструкторское бюро «Южное» им. М. К. Янгеля»,
Украина

² Днепропетровский национальный университет им. Олеся Гончара, Украина

ОЦЕНКА ВЛИЯНИЯ СЖАТИЯ СПУТНИКОВЫХ СНИМКОВ НА РЕЗУЛЬТАТЫ ИХ КЛАССИФИКАЦИИ

Рассчитаны статистические характеристики исходного многоспектрального спутникового снимка, а также снимков, сжатых наиболее распространенными алгоритмами сжатия графической информации с потерями JPEG, JPEG2000 и MRSID. Определены потери информационного сжатия спутниковых снимков и выполнена оценка влияния сжатия на результаты неконтролируемой классификации ISODATA путем количественного и качественного сравнения результатов классификации исходного и сжатых изображений. Выбран алгоритм сжатия графической информации, который вносит наименьшие погрешности в результаты классификации.

Ключевые слова: спутниковые снимки, алгоритмы JPEG, JPEG2000, MRSID, потери сжатия, неконтролируемая классификация, алгоритм ISODATA.

Введение

Среди различных видов космической деятельности дистанционное зондирование Земли (ДЗЗ) является одним из самых перспективных способов получения геопространственных данных, т.к. по сравнению с традиционными контактными способами обладает следующими достоинствами:

- оперативность и достоверность,
- безопасность и доступность;
- способность исследования огромных территорий в сжатые сроки;
- разнообразие решаемых задач и возможность применения результатов исследований в различных отраслях науки и народного хозяйства.

Благодаря этому ДЗЗ из космоса относится к одной из наиболее успешно и динамично развивающихся сфер современного информационного общества [1-3]. Научно-технические достижения последних лет в деле создания и развития космических систем, технологий обработки, хранения, интерпретации и использования получаемых данных многократно расширили круг и масштаб задач, решаемых с помощью ДЗЗ. Растет число потребителей: данные космических съемок активно используются не только в научных и производственных целях, но и в повседневной жизни людей.

1. Формулировка проблемы

Наиболее востребованными являются данные ДЗЗ высокого пространственного разрешения, особенно мультиспектральные. Такие снимки могут иметь размер более гигабайта и с целью экономии дискового пространства для них часто применяются алгоритмы сжатия, как без потерь (архиваторы ZIP, RAR, 7Z и др.) уменьшающими объем на несколько процентов, так и с потерями (JPEG, ECW, MRSID), сжимающими в десятки раз в зависимости от допустимых потерь. Особенно эффективно сжатие на борту спутника, которое позволяет существенно снизить требования к пропускной способности радиолинии сброса данных ДЗЗ, что крайне важно для микро- и наноспутников.

При сжатии с потерями очень важно правильно задавать коэффициент сжатия с учетом возможности проведения дальнейшей обработки снимка, в частности, классификации.

2. Актуальность исследований

Современные мультиспектральные (многозональные) съемочные системы формируют несколько отдельных изображений для нескольких спектральных зон, что затрудняет проведение визуальной классификации изображений, поэтому для них реализованы специальные алгоритмы классификации [4-6].

Известны два типа классификации:

- классификация с обучением (контролируемая классификация);
- классификация без обучения (неконтролируемая классификация).

Наиболее популярными среди алгоритмов неконтролируемой классификации (кластеризации) являются:

- алгоритм K-средних (K-Means);
- алгоритм ISODATA (Iterative Self-Organizing Data Analysis Technique).

Достоинство этих алгоритмов – возможность выполнения классификации спутниковых снимков в автоматическом режиме без обучения.

3. Цель и задачи исследований

Основной целью исследований является оценка влияния сжатия с потерями на результаты неконтролируемой классификации спутниковых снимков.

4. Методы исследований

Исследования проводились путем сравнения результатов классификации исходного и сжатых изображений.

Алгоритм кластеризации k-means (алгоритм k-внутригрупповых средних) основан на минимизации функционала Q суммарной выборочной дисперсии, который характеризует разброс элементов относительно центров кластеров:

$$Q = \sum_i |X_i| \sum_{x \in X_i} d(x, c_i) \rightarrow \min, \quad (1)$$

где $c_i = \frac{1}{|X_i|} \sum_{x \in X_i} x$ - центр кластера X_i .

При неконтролируемой классификации методом K-средних пользователь задает количество классов, затем вычисляются средние для классов и пиксели итеративно группируются в ближайший класс, используя технологию минимального расстояния. При каждой итерации заново считаются средние по классу, и по новому классифицируются пиксели по отношению к новым средним.

Алгоритм ISODATA относится к классу эвристических алгоритмов и использует минимальное спектральное расстояние для определения соответствующего кластера (класса) для каждого пикселя.

В него включены процедуры удаления, объединения и разделения кластеров. Процедура удаления кластеров выполняется, если число $|X_i|$ элементов в i-м кластере меньше заданного, т.е. при $|X_i| < L$. Элементы этого кластера распределяются

по другим кластерам, а его центр c_i удаляется из списка центров кластеров. Процедура разделения кластеров выполняется, если дисперсия i-го кластера D_i больше Q_s . В данном случае i-й кластер разделяется на два кластера.

Для разделения кластера вычисляются покомпонентные дисперсии:

$$D_{ik} = \frac{1}{|X_i|} \sum_{x_j \in X_i} \|x_{jk} - c_{jk}\|^2, \quad k=1, \dots, n, \quad (2)$$

где n - количество пикселей в i-м кластере.

Выбирается та l-я компонента, для которой $D_{il} > D_{is}$ для всех $l \neq s$ и осуществляется разделение i-го кластера по l-й компоненте. При этом пересчитываются новые центры кластеров.

Слияние кластеров. Если расстояние между двумя центрами кластеров достаточно мало, то эти кластеры объединяются в один кластер. Для реализации этой процедуры вычисляется расстояние между двумя центрами кластеров:

$$I_v = \|c_i - c_j\|, \quad \text{для всех } i \neq j. \quad (3)$$

Если $I_{ij} < Q_c$, то кластеры X_i и X_j следует объединить. Новый центр кластера вычисляется по формуле

$$c = \frac{c_i |X_i| + c_j |X_j|}{|X_i| + |X_j|}. \quad (4)$$

5. Исходные данные для исследований

Так как для Украины наиболее актуальной и массовой сферой применения спутниковых снимков является агромониторинг [7-9], то в качестве исходных данных использовался фрагмент многоспектрального снимка сельхозугодий (рис. 1) со спутника «Сич-2» в формате GEOTIFF (зеленый (G), красный (R) и ближний (NIR) ИК каналы). Гистограмма на рис. 1 показывает зависимость количества пикселей от их значений для каждого канала (диапазон значений 0 – 255 соответствует радиометрическому разрешению снимка 8 бит).

6. Порядок проведения исследований

Было выполнено сжатие с потерями наиболее распространенными методами: с использованием дискретного косинусного преобразования (формат JPEG) и с использованием дискретного вейвлет-преобразования [10-12] (форматы JPEG2000 и MRSID).

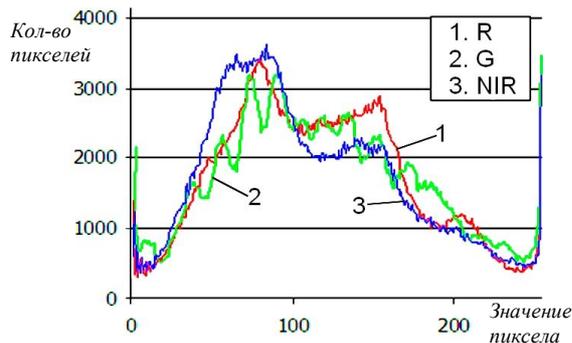


Рис. 1. Исходный снимок и его гистограммы каналов

Использовалось максимально возможное сжатие, при котором потери визуально незаметны (рис. 2-4 и таблица 1).

Как видно из рис. 2-4 и таблицы 1, формат MRSID обеспечивает наилучшее сжатие, но при этом вносит наибольшие потери, а формат JPEG2000 вносит наименьшие потери при среднем сжатии.

На рис. 5 приведены основные статистические характеристики изображений, рассчитанные для каждого метода сжатия.

На рис. 6-8 показаны компоненты исходного изображения, удаленные в процессе сжатия (потери сжатия), а также основные статистические характеристики, рассчитанные для каждого метода сжатия (статистические моменты и гистограммы для каждого канала).

Из рис. 6-8 следует, что сжатия JPEG и MRSID удаляют преимущественно сигнальные компоненты снимка, а сжатие JPEG2000 удаляет преимущественно шумовые компоненты.

Затем была выполнена неконтролируемая классификация исходного снимка (количество классов задавалось равным 5) алгоритмам k-means (рис. 9) и ISODATA (рис. 10), которые для данного фрагмента дали близкие результаты.

Поэтому далее проводилось сравнение результатов классификации исходного и сжатого снимков только для алгоритма ISODATA (рис. 11-13).

На рис. 14 показаны погрешности классификации, обусловленные потерями сжатия – процент пикселей, которые не совпали с результатами классификации исходного снимка.

Похожие результаты были получены и на других трехканальных снимках.

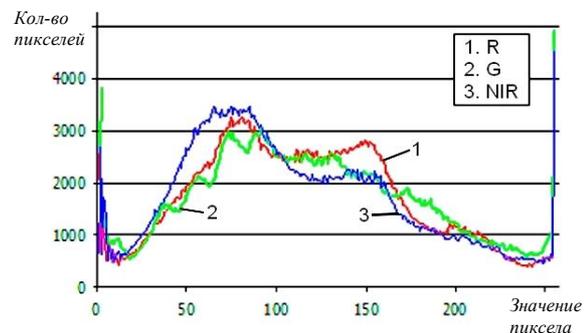


Рис. 2. Снимок после сжатия JPEG и его гистограммы каналов

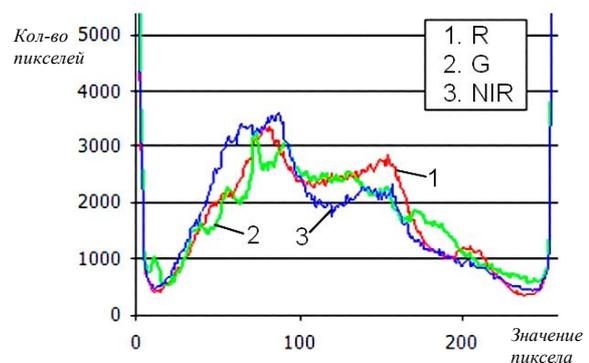


Рис. 3. Снимок после сжатия JPEG2000 и его гистограммы каналов

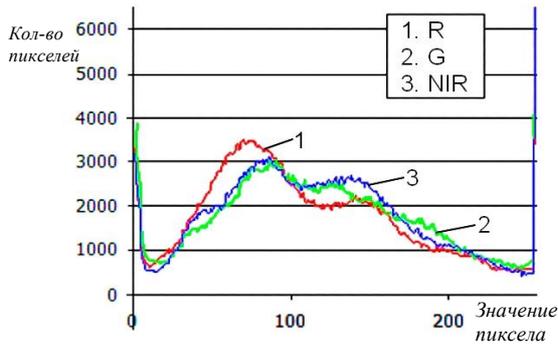
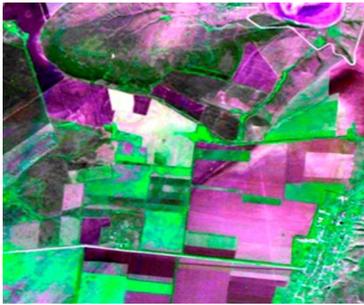


Рис. 4. Снимок после сжатия MRSID и его гистограммы каналов

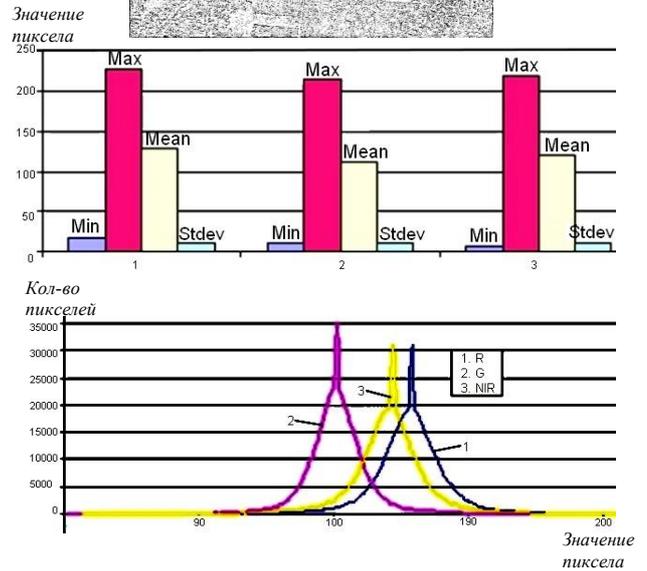
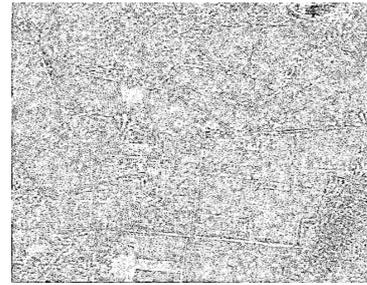


Рис. 6. Потери сжатия JPEG

Таблица 1

Статистические характеристики исходного и сжатых снимков

Формат файла	Размер файла	Коэф. сжатия	Математическое ожидание			Дисперсия		
			R	G	NIR	R	G	NIR
tif	1415400	1	111,03	112,09	104,46	63,72	66,54	64,03
jpg	88890	15,9	111,33	112,11	104,76	63,05	66,30	63,57
jp2	83738	16,9	111,05	112,14	104,55	62,11	66,02	62,60
sid	61757	22,9	105,09	112,87	111,91	63,93	66,49	63,46

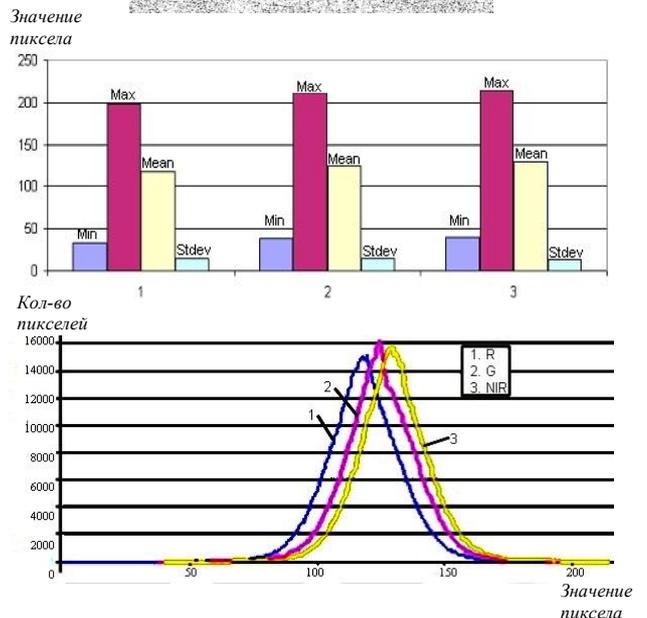
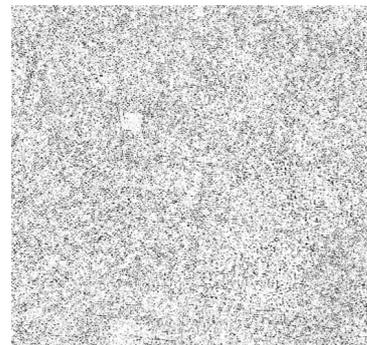


Рис. 7. Потери сжатия JPEG2000



Рис. 5. Статистические характеристики изображений: 1 – исходное изображение, 2 – сжатие JPEG, 3 – сжатие JPEG2000, 4 – сжатие MRSID

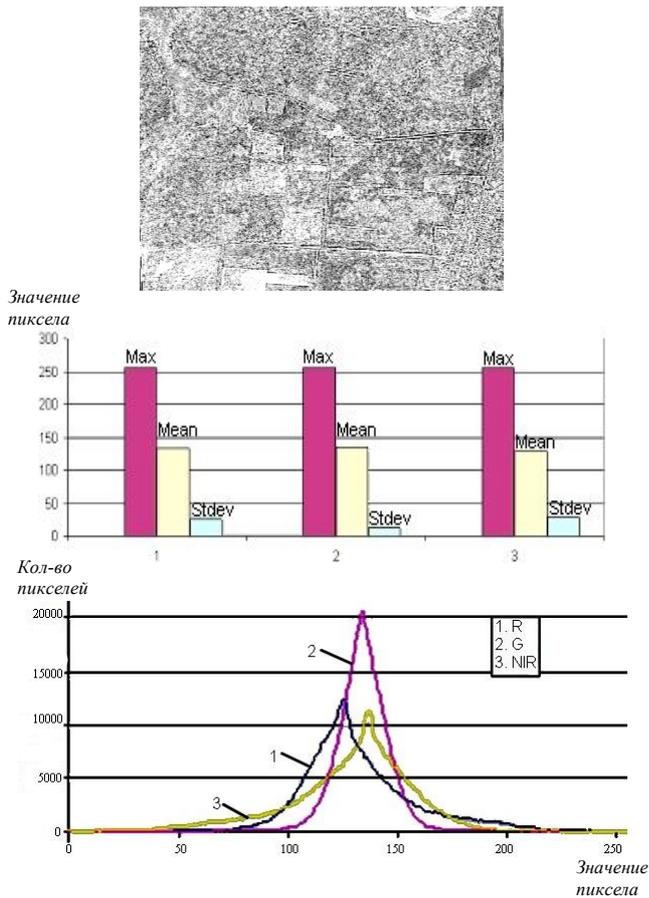


Рис. 8. Потери сжатия MRSID

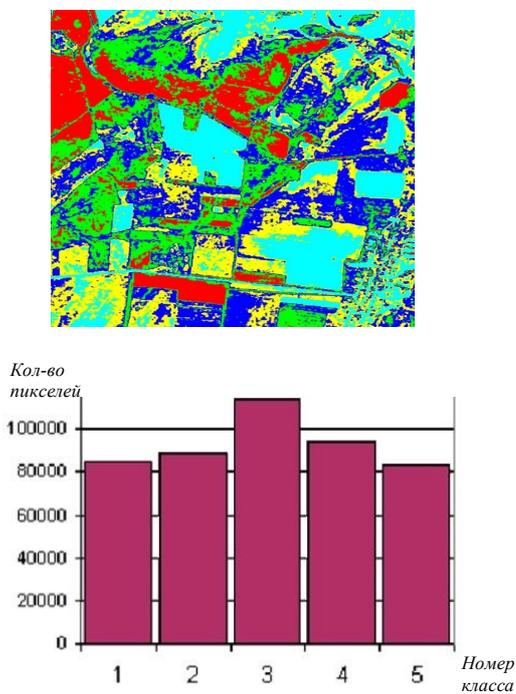


Рис. 9. Алгоритм k-means и его гистограмма классов

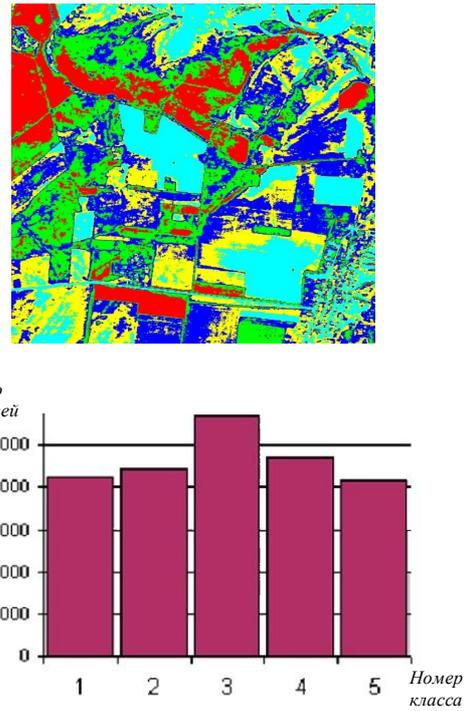


Рис. 10. Алгоритм ISODATA и его гистограмма классов

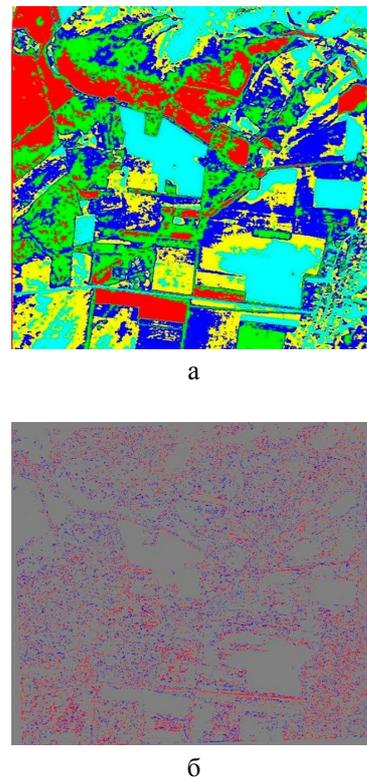
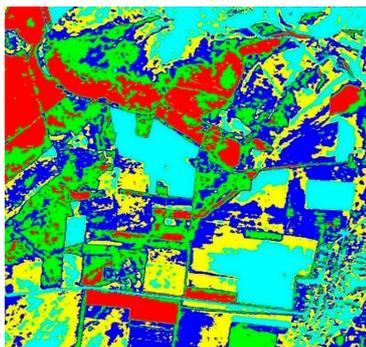
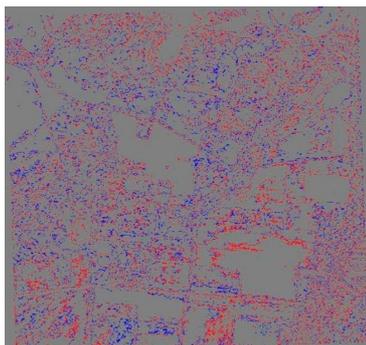


Рис. 11. Сжатие JPEG: а – результат классификации, б – отличие от исходного



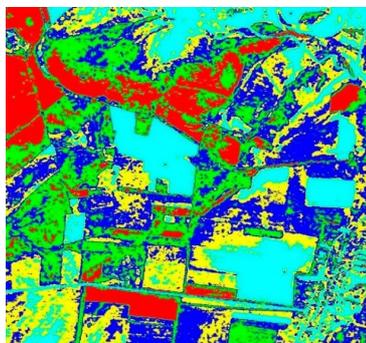
а



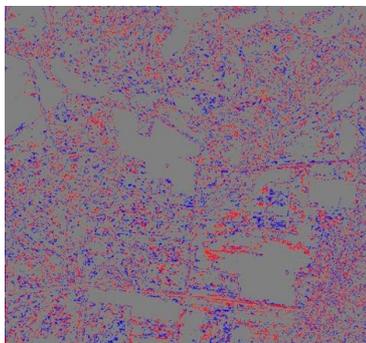
б

Рис. 12. Сжатие JP2:

а – результат классификации,
б – отличие от исходного



а



б

Рис. 13. Сжатие MRSID:

а – результат классификации
б – отличие от исходного

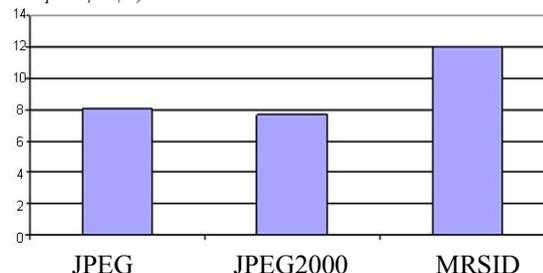
Погрешности
классификации (%)

Рис. 14. Погрешности классификации (%)

Заключение

Алгоритмы неконтролируемой классификации K-Means и ISODATA при небольшом (5...7) количестве классов и спектральных каналов (3...4) дают близкие результаты.

Формат MRSID обеспечивает наилучшее сжатие, но при этом вносит наибольшие погрешности в результаты классификации (около 12%). Формат JPEG2000 вносит наименьшие погрешности в результаты классификации (менее 8%) при среднем сжатии (около 17).

Литература

1. Мозговой, Д. К. Обработка спутниковых снимков при решении прикладных задач [Текст] / Д. К. Мозговой // Международный научно-практический форум «Наука и бизнес»: тез. докл., 29–30 июня 2015 г. – Днепропетровск, 2015. – С. 191-194.
2. Мозговой, Д. К. Использование данных MODIS и ASTER для решения актуальных прикладных задач [Текст] / Д. К. Мозговой // Международная научно-практическая конференция «Стратегические решения информационного развития экономики и общества»: тез. докл., 17–19 июля 2013 г. – п. Научный, Крым, 2013. – С. 113–114.
3. Remote Sensing and GIS Application for Environmental Monitoring and Accidents Control in Ukraine [Text] / D. K. Mozgoviy, O. I. Parshina, V. I. Voloshin, Y. I. Bushuev // Geographic Uncertainty in Environmental Security; Edited by A. Morris, S. Kokhan. – Dordrecht: Springer, NATO Public Diplomacy Division, 2007. – P. 259-270.
4. Мозговой, Д. К. Использование многоспектральных снимков для классификации посевов сельскохозяйственных культур [Текст] / Д. К. Мозговой, О. В. Кравец // Экологія та ноосферологія. – 2009. – Т. 20, № 1–2. – С. 54-58.

5. Кравец, О. В. Классификация сельхозкультур по разновременным многоспектральным снимкам [Текст] / О. В. Кравец, Д. К. Мозговой // X Міжнародна молодіжна науково-практична конференція «Людина і космос» : зб. тез., 9–11 квітня 2008 р / Національний центр аерокосмічної освіти молоді України ім. О. М. Макарова. – Дніпропетровськ, 2008. – С. 501.

6. Мозговой, Д. К. Классификация малоразмерных объектов на спутниковых снимках [Текст] / Д. К. Мозговой, О. В. Кравец // Экологія та ноосферологія. – 2009. – Т. 20, № 3–4. – С. 26–30.

7. Мозговой, Д. К. Информационная система "Агрокосмос" [Текст] / Д. К. Мозговой, В. И. Волошин // VII українська конференція з космічних досліджень : тез. докл., Євпаторія, НЦУВКЗ, 3–8 вересня 2007 р. – К. : ІКІ НАНУ, НКАУ, 2007. – С. 166.

8. Мозговой, Д. К. Геоінформаційне забезпечення агросфери – система "Агрокосмос" [Текст] / Д. К. Мозговой, В. І. Волошин, А. Г. Тараріко // Міжнародна науково-практична конференція «Сучасні проблеми створення і ефективного використання єдиного геоінформаційного простору України» : тез. докл., 12–13 грудня 2007 р. / Інститут проблем національної безпеки. – Київ, 2007. – С. 65–66.

9. Dolinets, Y. Specialists training in ERS [Text] / Y. Dolinets, D. Mozgovoy // Advanced Space Technolo-

gies For The Humankind Prosperity. International Conference, 18–20 April 2007. – Dnepropetrovsk : Yuzhnoe State Design Office, 2007. – P. 123.

10. Мозговой, Д. К. Використання вейвлетів в обробці супутникових знімків високого просторового розрізнення [Текст] / Д. К. Мозговой, В. М. Корчинський, О. В. Кравець // Вісник ДНУ. Ракетно-космічна техніка. – 2009. – Вип. 13, т. 1, № 17/4. – С. 65–71.

11. Korchinskij, V. M Wavelet-technologies of digital signals transform while remote sensing data processing [Text] / V. M. Korchinskij, D. K. Mozgovoy, O. V Kravets // Advanced Space Technologies For The Humankind Prosperity. International Conference, 15–17 April 2009. – Dnepropetrovsk : Yuzhnoe State Design Office, 2009. – P. 103.

12. Kravets, O. V. Wavelet transform in remote sensing data processing [Text] / O. V. Kravets, D. K. Mozgovoy // Engineer of the 3rd Millennium. Students Scientific conference / National University of Railway Transport, Dnepropetrovsk, 12 May 2009. – Dnepropetrovsk, 2009. – P. 49–50.

13. Richards, John A. Remote Sensing Digital Image Analysis [Text] / John A. Richards, Xiuping Jia. – Springer-Verlag Berlin Heidelberg, 2006. – P. 251.

14. Landgrebe, David A. Signal Theory Methods In Multispectral Remote Sensing [Text] / David A. Landgrebe. – New Jersey : John Wiley & Sons, Inc., Hoboken, 2003. – P. 163.

Поступила в редакцію 28.10.2015, рассмотрена на редколлегии 18.11.2015

ОЦІНКА ВПЛИВУ СТИСКУ СУПУТНИКОВИХ ЗНІМКІВ НА РЕЗУЛЬТАТИ ЇХНЬОЇ КЛАСИФІКАЦІЇ

**В. М. Маслей, К. Г. Білоусов, Д. К. Мозговой, В. С. Хорошилов,
О. С. Бушанська, Д. В. Маслей, Д. О. Хмара**

Розраховано статистичні характеристики вихідного багатоспектрального супутникового знімка, а також знімків, стислих найбільш поширеними алгоритмами стиснення графічної інформації з втратами JPEG, JPEG2000 і MrSid. Визначено втрати інформаційного стиснення супутникових знімків і виконано оцінку впливу стиснення на результати неконтрольованої класифікації ISODATA шляхом кількісного та якісного порівняння результатів класифікації вихідного і стиснутих зображень. Обрано алгоритм стиснення графічної інформації, який вносить найменші похибки в результати класифікації.

Ключові слова: супутникові знімки, алгоритми JPEG, JPEG2000, MRSID, втрати стиску, неконтрольована класифікація, алгоритм ISODATA.

**ESTIMATION OF THE LOSSY COMPRESSION OF SATELLITE PICTURE INFLUENCE
ON THE CLASSIFICATIONS RESULTS**

*V. M. Masley, K. G. Belousov, D. K. Mozgovoj, A. L. Khoroshilov,
O. S. Bushanska, D. V. Masley, D. A. Khmara*

Calculated the statistical characteristics of the original multispectral satellite imagery as well as imagery, the most common compression algorithms compress graphics lossy JPEG, JPEG2000 and MRSID. Loss of information defined compression of satellite images and estimated impact of compression on the results of unsupervised classification ISODATA by quantitative and qualitative comparison of the results of the classification of the original and compressed images. Compression algorithm of graphic information was selected, which makes the smallest error in the classification results.

Keywords: satellite picture, JPEG, JPEG2000, MRSID algorithms, loss of the compression, uncontrolled classification, ISODATA algorithm.

Маслей Владимир Никитович – Главный конструктор и начальник, Проектно-Конструкторского бюро космических аппаратов, систем измерений и телекоммуникаций, Государственное предприятие «Конструкторское бюро «Южное» им. М. К. Янгеля», Днепропетровск, Украина.

Белузов Константин Георгиевич – Заместитель Главного конструктора и начальника Проектно-Конструкторского бюро космических аппаратов, систем измерений и телекоммуникаций, Государственное предприятие «Конструкторское бюро «Южное» им. М. К. Янгеля», Днепропетровск, Украина.

Мозговой Дмитрий Константинович – канд. техн. наук, доцент факультета физики, электроники и компьютерных систем, Днепропетровский национальный университет им. Олесь Гончара, Днепропетровск, Украина.

Хорошилов Виктор Сергеевич – д-р техн. наук, проф., Заслуженный деятель науки и техники Украины, начальник расчетно-теоретического сектора Проектно-Конструкторского бюро космических аппаратов, систем измерений и телекоммуникаций, Государственное предприятие «Конструкторское бюро «Южное» им. М. К. Янгеля», Днепропетровск, Украина.

Бушанская Александра Сергеевна – аспирант факультета физики, электроники и компьютерных систем, Днепропетровский национальный университет им. Олесь Гончара, Днепропетровск, Украина.

Маслей Дмитрий Владимирович – соискатель кафедры электронных средств телекоммуникаций, Днепропетровский национальный университет им. Олесь Гончара.

Хмара Денис Александрович – инженер отдела Проектно-Конструкторского бюро космических аппаратов, систем измерений и телекоммуникаций, Государственное предприятие «Конструкторское бюро «Южное» им. М. К. Янгеля», Днепропетровск, Украина.