

Устранение информационной избыточности гиперспектральных растровых изображений методом «хорошо приспособленного» базиса

Д.Ю. Васин¹, В.П. Громов¹, П.А. Пахомов¹

¹Нижегородский государственный Национальный исследовательский университет им. Н.И. Лобачевского, Гагарина 23, Нижний Новгород, Россия, 603952

Аннотация. Работа посвящена рассмотрению вопросов устранения информационной избыточности гиперспектральных растровых данных (ГСИ). Перечислены характерные свойства ГСИ, приводится краткая характеристика существующих методов компрессии ГСИ. Рассматривается возможность применения локальных, однородных «хорошо приспособленных» базисных функций (ЛОХПБФ) для устранения информационной избыточности и адаптивного сжатия ГСИ. Предложен алгоритм построения системы ЛОХПБФ для ГСИ, основанной на чебышевском приближении. Приведены результаты вычислительных экспериментов, в том числе с использованием графического процессора. Показана эффективность предлагаемого метода адаптивного сжатия ГСИ.

1. Введение

В настоящее время средства дистанционного зондирования Земли (ДЗЗ) авиационного и космического базирования являются достаточно эффективными в решении различных задач мониторинга поверхности Земли [1]. Большинство современных систем являются многоканальными, где под многоканальным режимом в широком смысле понимается формирование изображений одного и того же участка поверхности с использованием нескольких частот (длин волн), поляризаций, углов визирования и т.д. Одним из перспективных типов многоканальных систем ДЗЗ являются средства гиперспектральной съемки, которые «перекрывают» оптический и ближний инфракрасный диапазоны электромагнитных волн с высоким спектральным разрешением (порядка единиц нанометров) и хорошим пространственным разрешением (от единиц до десятков метров) [1], формируя одновременно сотни практически совмещенных между собой изображений. Примерами таких систем и гиперспектральных сенсоров являются AVIRIS, HYDICE, Hyperion, CASI, CHRIS-PROBA и др.

Получаемые гиперспектральные данные (ГСД) используются для решения широкого круга разнообразных прикладных тематических задач [1, 2].

При передаче ГСД на Землю и пользователям, при обработке гиперспектральных изображений (ГСИ) и извлечении из них полезной информации, а также хранении возникает ряд проблем [1–3], при этом, одной из центральных является огромный объем информационных данных, которые необходимо передавать по каналам связи и обрабатывать [1, 3–7]. Поэтому актуальна задача развития существующих и поиска новых методов сжатия ГСИ.

В монографии [8] приведен достаточно полный обзор существующих методов компрессии

изображений в системах оперативного ДЗЗ, отражена специфика задачи обработки изображений в таких системах и вытекающие из нее требования к методам бортовой обработки получаемых видеоданных. Среди них выделим задачи обеспечения высокой степени сжатия данных и низкой вычислительной и структурной сложности используемых методов компрессии. В то же время, проблема сжатия ГСИ актуальна при организации хранилищ растровых данных ГСИ и их передаче по каналам связи между различными пользователями.

При кодировании больших объемов экспериментальных данных широко используются методы на основе дисперсионного и факторного анализов. При этом находятся базисные функции, в некотором смысле приспособленные к кодированию рассматриваемых данных. Получаемые по методу ортогональных компонент базисные функции оптимальны в смысле среднестатистической среднеквадратической погрешности, а при предварительной нормировке кодируемых данных по длительности и энергии оптимальны в смысле минимума коэффициентов разложения [17].

Однако, с практической точки зрения эти методы достаточно трудоемки в вычислительном плане и требуют определенных ресурсов по памяти, поскольку необходимо вычислять собственные вектора матриц ковариации, полученных по набору ГСИ.

В Нижегородском госуниверситете в научной школе засл. деят. науки РФ, д.т.н., проф. Васина Ю.Г. (1940 – 2017), более 40 лет ведутся исследовательские работы по созданию программно-аппаратных комплексов обработки большеформатной сложноструктурированной графической информации. Без сомнения, растровые данные ДЗЗ, в частности ГСИ, напрямую относятся к данному виду графической информации.

В работе [18] был предложен контекстный метод сжатия без потерь RGB- и мультиспектральных изображений. Данная работа является продолжением исследований авторов по поиску квазиоптимальных систем кодирования ГСИ, реализующих сжатие с потерями, но обеспечивающих несколько большие коэффициенты сжатия по сравнению с методами сжатия без потерь.

2. Характерные свойства ГСИ ДЗЗ

ГСИ ДЗЗ характеризуются рядом специфических свойств, отличающих их от обычных черно-белых и цветных оптических изображений:

1. ГСИ обладают высокой степенью избыточности, поскольку изображения в двух соседних спектральных каналах обычно имеют коэффициент корреляции, близкий к единице [1, 4, 15], в то время, как коэффициенты корреляции цветовых компонент R, G и B цветных изображений обычно порядка 0,8, то есть заметно меньше [16]. Такое отличие связано с тем, что для соседних каналов ГСИ отличия длин волн гораздо меньше, чем длин волн при RGB представлении цветных оптических изображений;

2. для RGB цветных изображений отношения сигнал-шум (ОСШ) в компонентных изображениях обычно полагают идентичными или достаточно близкими; [16]. Для компонент ГСИ ОСШ обычно существенно различаются. Из практики, для относительно небольшого числа спектральных каналов ОСШ находится в пределах единиц дБ, но для большинства спектральных каналов ОСШ порядка 20...30 дБ. Это связано, прежде всего, с существенным различием динамических диапазонов данных и с отличием уровня помех в разных спектральных каналах[1, 13];

3. динамический диапазон для каждого n -го спектрального канала можно характеризовать разностью максимального $P^{max}(n)$ и минимального $P^{min}(n)$ значений. Значения $P^{max}(n)$ могут варьироваться от нескольких десятков до десятков тысяч, а значения $P^{min}(n)$ имеют порядок единиц и десятков, причем они могут быть как положительными, так и отрицательными в связи с особенностями используемых сенсоров. Поэтому для представления ГСИ ДЗЗ обычно используют 2 байта на отсчет, где один из битов – знаковый.

3. Краткая характеристика существующих методов компрессии ГСИ

Все существующие методы компрессии изображений в целом, разделяются на две группы, обеспечивающих сжатие без потерь, и с потерями.

Сжатие без потерь обеспечивает КС ГСИ порядка 3 – 4 [4] при учете межканальной корреляции

и около 1.5 – 2 при сжатии каждого изображения отдельно. Несомненное преимущество данной группы методов состоит в том, что при их использовании в ГСИ не вносятся никаких искажений. Вместе с тем, получаемые значения КС не всегда удовлетворяют требованиям практики. Поэтому актуальна задача разработки и применения методов сжатия ГСИ с потерями [5–7], которые способны обеспечивать несколько большие значения КС по сравнению с методами первой группы. Для возможности применения методов данной группы исходные растровые данные ГСИ представим в виде векторов, получаемых в результате последовательного, построчного сканирования исходных растров ГСИ.

В основе этих методов лежат идеи разложения исходных сигналов по той или иной системе базисных функций (СБФ) с заданной точностью приближения ε . При этом, проблема оптимального кодирования ГСИ сводится к поиску такой СБФ, которая при заданной среднеквадратичной ошибке δ обеспечивает минимальное, либо близкое к нему, число таких функций $\varphi^1(t), \varphi^2(t), \dots, \varphi^m(t)$. Тогда процесс $f(t)$ ($t_1 \leq t \leq t_2$) приближенно представим в виде:

$$\tilde{f}(t) = \sum_{k=1}^m C_k \varphi^k(t) \text{ базисных функций } \varphi^1(t), \varphi^2(t), \dots, \varphi^m(t). \text{ Коэффициенты } C_1, C_2, \dots, C_m \text{ при}$$

этом рассматривается как код кривой $f(t)$. Погрешность приближения при этом:

$$\varepsilon(t) = f(t) - \sum_{k=1}^m C_k \varphi^k(t). \text{ Очевидно, что для различных видов получаемых ГСД потребуются и}$$

различные оптимальные СБФ.

При разработке методов сжатия ГСД с потерями возникает ряд практических моментов:

- целью получения ГСД является решение конечных задач распознавания и классификации (классификации типов земных покровов, обнаружение и локализация загрязненных участков местности, и т.д.). При этом часто нет необходимости обрабатывать данные во всех каналах ГСД, достаточно использовать весьма ограниченный набор наиболее информативных, в смысле решения данной тематической задачи, каналов [2]. В каждом конкретном случае набор этих каналов может быть различным, но однозначно, что пользователю необходимо передавать не полный массив ГСД, а лишь его часть (обычно не более 8...12 каналов), что уже эквивалентно сжатию.

- применение сжатия с потерями не всегда безусловно приводит к ухудшению эффективности решения конечных задач. При определенных условиях эффективность решения конечных задач на сжатых с потерями ГСД может оставаться практически такой же, как и при использовании исходных (несжатых или сжатых без потерь) ГСД. [8, 9]

- для методов сжатия ГСИ с потерями в настоящее время не существует общепринятых метрик качества [12-14].

Отметим, что основной шаг при аппроксимации исходного сигнала – выбор информативной СБФ ведется полуэмпирически с использованием опыта и интуиции исследователя, либо исходя из внешнего сходства исходного сигнала и СБФ. Это может приводить к тому, что оптимальная, либо близкая к ней, СБФ может быть не найдена, особенно в случае сложной формы первоначального сигнала.

В этом плане более интересны объективные методы выбора СБФ, в которых не опираются на знания человека, а исходят лишь из свойств совокупности исходных непрерывных данных. Одной из таких оптимальных СБФ является система собственных векторов матрицы ковариации, вычисленной для заданного множества исходных данных, а соответствующие им собственные значения характеризуют точность приближения [11]. Однако, данный метод достаточно громоздкий в вычислительном плане, кроме того, получаемая таким способом СБФ, позволяет с заданной точностью кодировать не каждую функцию исходного набора сигналов, а только в среднем по множеству. При этом основная часть данных будет кодироваться с заданной точностью приближения, однако кодирование экстремальных данных может происходить со значительной ошибкой. Очевидно, что при кодировании ГСИ это является нежелательным свойством получаемой СБФ.

С целью уменьшения вычислительной сложности алгоритма формирования оптимальной СБФ,

был предложен алгоритм построения системы «хорошо приспособленных» базисных функций (СХПБФ) [9].

На конечном интервале обработки исходные дискретные данные $x^j = \{x_1^j, x_2^j, \dots, x_n^j\}$, $j=1, 2, \dots, N$ можно представить как точки в n -мерном пространстве R_n , с координатами $\{x_1, x_2, \dots, x_n\}$. Тогда некоторому набору экспериментальных кривых будет соответствовать множество X точек $x^j = \{x_1^j, x_2^j, \dots, x_n^j\}$, $j=1, 2, \dots, N$ в пространстве R_n [9].

При такой геометрической интерпретации разработан алгоритм построения СХПБФ, основанный на экспериментально обнаруженном, но достаточно общем, факте сравнительно малой размерности множеств, отвечающих первоначальным описаниям экспериментальных данных.

Если исходить из того факта, что множество X в пространстве R_n имеет малую размерность, то аппроксимируем множество X некоторым линейным многообразием L_m размерности $m < n$ и выберем ортонормированный базис линейного многообразия L_m за СХПБФ для кодирования точек множества X . Тогда задача сокращения избыточности в исходном описании экспериментальных данных может быть сформулирована следующим образом [10].

В n -мерном пространстве R_n с координатами $\{x_1, x_2, \dots, x_n\}$ задано множество точек $x^j \in X$, $j=1, 2, \dots, N$. Требуется найти ортонормированный базис u_1, u_2, \dots, u_m некоторого подпространства R_m , в котором с заданной точностью ε , в смысле некоторой заданной метрики ρ , лежат все точки $x^j \in X$, $j=1, 2, \dots, N$. Как только указанный базис найден, любая точка $x \in X$ может быть представлена с заданной точностью ε при помощи m новых признаков C_k ,

вычисляемых согласно формуле $C_k = \sum_{i=1}^n x_i u_i^k$, $k=1, 2, \dots, m$. Ординаты декодированной точки:

$$\tilde{x}_i = \sum_{k=1}^m C_k u_i^k, \quad k=1, 2, \dots, m$$

Изначально, кодирование по СХПБФ, основанной на чебышевских приближениях, было предложено в задачах обработки медицинских кривых (ЭКГ, энцефалограмм и др.) [9, 10].

Заметим также, что при кодировании обычно важна не столько возможно большая точность приближения ε с заданным числом m базисных векторов, сколько минимизация самого числа базисных векторов m при заданной точности приближения ε . Доказано, что при заданной точности ε , в предлагаемый хорошо приспособленный базис не входит самое большее три последних орта по сравнению с оптимальным кодированием в среднем [10].

Рассматриваемый в статье способ сжатия, основанный на чебышевских приближениях, обеспечивает ошибку кодирования, не большую заданной для всех точек выпуклой оболочки исходного множества векторов, в то время как методы кодирования, опирающиеся на метод главных компонент, обеспечивают достаточно малую ошибку в среднем по всему исходному множеству.

Рассматриваемый метод кодирования является квазиоптимальным, его достоинство заключается в том, что он допускает достаточно простую практическую реализацию для больших размерностей исходных данных.

Это и послужило отправной точкой для исследования возможности сжатия ГСИ по СХПБФ, основанной на чебышевских приближениях.

4. Алгоритм построения СХПБФ, основанной на чебышевских приближениях

Пусть $x^j \in X$, $j=1, 2, \dots, N$ – множество исходных точек, N – мощность множества X , δ – требуемая точность приближения (в %) и максимально допустимое количество L^{\max} базисных ортов СХПБФ ($L^{\max} \leq P$), L – число сформированных ортов СХПБФ после выполнения очередной итерации алгоритма, N_L – число точек множества X , которые к данному моменту уже

разложены с заданной точностью ε по уже сформированным базисным ортам U^L , \bar{N} – число непомеченных точек множества, участвующих в формировании очередного базисного орта.

Шаг 1. Инициализируем счетчики: $L=0$, $N_L=0$, $\bar{N} = N - N_L$.

Шаг 2. Линейно просматриваем \bar{N} непомеченных точек $x^j \in X$ и находим точку $x^{jL} \in X$ такую, что $\lambda_L = \max_j \|x^j\|$, $j = 1, 2, \dots, \bar{N}$.

Шаг 3. Если $L = 0$, то формируем ошибку разложения $\varepsilon = \delta * \lambda_0$, иначе **Шаг 4**.

Шаг 4. Увеличиваем счетчик L на единицу и находим очередной базисный орт U^L :

$$U^L = \frac{x^{jL}}{\lambda_L}$$

Шаг 5. Линейно просматриваем все непомеченные точки $x^j \in X$ и выполняем их кодирование по только что построенному орту U^L , т.е. для каждой такой точки получаем значение коэффициента $C^j = \sum_{i=1}^n (x_i^j, u_i^L)$, $j = 1, 2, \dots, \bar{N}$ и восстанавливаем координаты точек

x^j по построенному орту U^L : $\tilde{x}_i^j = C^j u_i^L$, $i = 1, 2, \dots, n$, $j = 1, 2, \dots, \bar{N}$.

Шаг 6. Корректируем метрику точек: $x_i^j = x_i^j - \tilde{x}_i^j$, $i = 1, 2, \dots, n$, $j = 1, 2, \dots, \bar{N}$.

Шаг 7. Линейно просматриваем все непомеченные точки $x^j \in X$ и помечаем те из них, для которых выполняется условие $\|x^j\| < \varepsilon$, $j = 1, 2, \dots, \bar{N}$; корректируем значение счетчика N_L , увеличивая его на число точек K , помеченных на данном шаге (разложенных по текущему орту U^L с требуемой точностью ε), $N_L = N_L + K$, и счетчика $\bar{N} = N - N_L$.

Шаг 8. Если $\bar{N} = 0$, т.е. в наборе не осталось непомеченных точек $x^j \in X$, $j = 1, 2, \dots, \bar{N}$ или число построенных ортов $L > L^{\max}$, то **КОНЕЦ АЛГОРИТМА**, иначе переход на **Шаг 2**. Исключение на каждом последующем шаге алгоритма тех точек x^j , которые к данному моменту уже представляются с заданной точностью ε , позволила существенно снизить вычислительную сложность алгоритма, что особенно важно на последних итерациях, когда производится достаточно много вычислений.

5. Результаты практических экспериментов

Практические эксперименты по кодированию ГСИ ДЗЗ проводились на выборке из 360 полутоновых растровых изображений (данных спектральных каналов) с геометрическими размерами каждого растра $\sim 10^3 \times 10^3$ пикселей и глубиной цвета 16 бит/пикс. Растровое изображение каждого канала путем построчного сканирования превращалось в одномерный сигнал, размерности $\sim 10^6$ отсчетов. Таким образом, на входе алгоритма построения СХПБФ размерность исходного пространства $R_n \sim 10^6$, а мощность множества исходных векторов V : $H(V) = 360$.

Для каждого n – спектрального канала экспериментально определялся его динамический диапазон: $D(n) = 10 \lg \frac{P^{\max}(n) + 1}{P^{\min}(n) + 1}$ (дБ), где $P^{\max}(n)$ и $P^{\min}(n)$ максимальное и минимальное

значения яркости пикселей в канале n , $n=1, 2, \dots, H(V)$. Минимальное и максимальное значение яркостей по всем каналам: $D^{\min} = 4 \text{ дБ}$, $D^{\max} = 24 \text{ дБ}$. Полученный интервал

$I = D^{\max} - D^{\min} = 20 \text{ дБ}$ был равномерно разбит на 10 групп. Распределение каналов по группам приведено в таблице 1 и на рисунке 1.

Далее для каждой группы каналов строилась СХПБФ для разных значений требуемой точности приближения δ , а коэффициент сжатия $K_{сж}(\delta)$ определялся для каждого значения δ по формуле

$$K_{сж}(\delta) = \left(1 - \frac{N_{орт}(\delta)}{P}\right) * 100\%, \text{ где } N_{орт}(\delta) - \text{число сформированных ортов СХПБФ по}$$

большим динамическим диапазоном внесенные искажения будут незаметны. Разбиение на группы нивелирует различие каналов по динамическому диапазону.

Таблица 1. Распределение каналов по группам в зависимости от динамического диапазона.

Группа	Динамический диапазон, дБ	Количество каналов	Номера каналов
1	4–6	2	4, 5
2	6–8	0	–
3	8–10	42	3, 6, 16 – 27, 38 – 65
4	10–12	5	2, 34 – 37
5	12–14	11	1, 7, 12 – 15, 28 – 30, 66, 67
6	14–16	4	31 – 33, 68
7	16–18	4	8 – 11
8	18–20	40	103, 199 – 203, 256 – 282, 329, 331 – 335, 338
9	20–22	126	69 – 71, 185 – 198, 204 – 255, 283 – 328, 330, 336, 337, 339 – 346
10	22–24	126	72 – 102, 104 – 184, 347 – 360



Рисунок 1. Гистограмма распределения числа спектральных каналов от номера группы.

Разбиение на группы представляется необходимым по следующим соображениям: в случае построения СХПБФ по исходным данным с принципиально различными динамическими диапазонами каналов $D(n)$, ошибка разложения ϵ не способна адекватно описывать искажения, вносимые при сжатии с потерями изображения в n -м спектральном канале.

Одно и то же значение ϵ в канале с малым $D(n)$ приведет к явной деградации качества изображения, а в канале с всем группам для фиксированного значения δ . Зависимость суммарного Ксж от точности приближения δ приведена на рисунке 2. Фактические значения в таблице 2.

Рисунок 3 содержит фрагменты изображений 107-го частотного канала и гистограммы распределения яркостей пикселей в нем. Рисунок 3а – исходное изображение, рисунок 3б –

восстановленное изображение с $\delta= 3\%$ по 94 ортам СХПБФ, рисунок 3с – ошибка восстановления с $\delta= 3\%$ по 94 ортам СХПБФ, рисунок 3d – восстановленное изображение с $\delta= 7\%$ по 30 ортам СХПБФ, рисунок 3е – ошибка восстановления с $\delta= 7\%$ по 30 ортам СХПБФ.

Таблица 2. Зависимость числа сформированных ортов ХПБФ от номера группы и точности приближения.

Группа	Точность приближения δ (%)						
	1	1,5	2	3	5	7	15
Число сформированных ортов СХПБФ							
1	2	2	2	2	2	2	2
2	–	–	–	–	–	–	–
3	42	42	42	31	8	3	2
4	5	5	5	4	3	2	1
5	11	11	11	10	6	5	1
6	4	4	4	4	3	2	1
7	4	4	4	4	4	4	1
8	37	31	25	10	4	3	3
9	126	99	69	16	6	5	2
10	121	80	43	13	7	4	3
Σ	352	278	205	94	43	30	16
$K_{сж}(\delta)$ (в %)	2	23	43	74	88	92	96

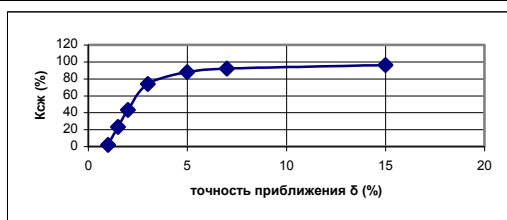


Рисунок 2. Зависимость $K_{сж}$ от точности приближения δ (%).

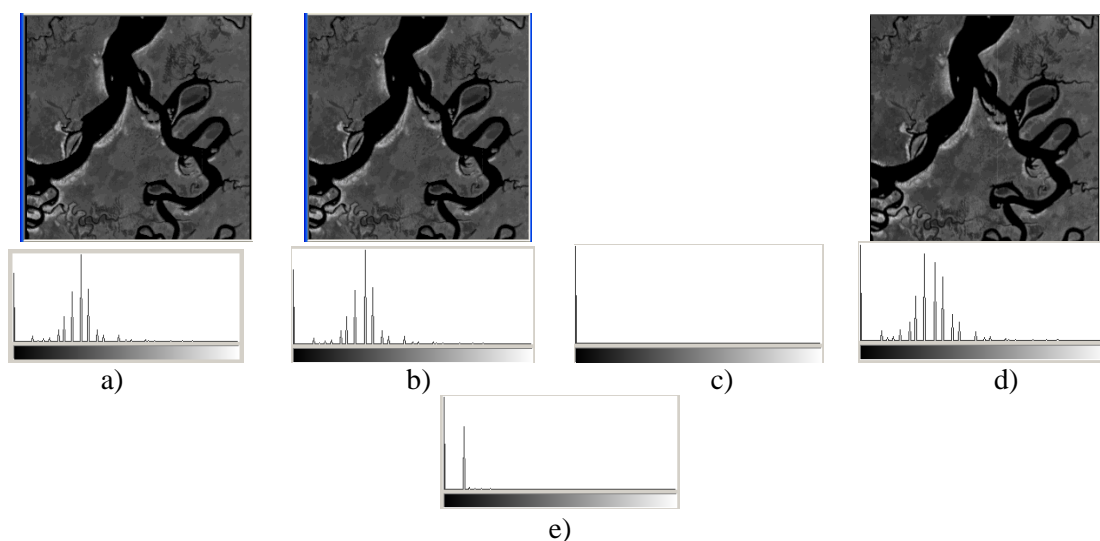


Рисунок 3. а) фрагмент исходного изображения 107 частотного канала; б) восстановленное изображение 107 частотного канала с $\delta= 3\%$ по 94 ортам СХПБФ; в) структура ошибки восстановления; г) восстановленное изображение 107 частотного канала с $\delta= 7\%$ по 30 ортам СХПБФ; д) структура ошибки восстановления.

Как видно из рисунка 3с, восстановление изображения канала по 94 ортам не содержит никаких искажений по сравнению с исходным, на рисунке 3е результат декодирования исходного сигнала по 30 ортам СХПБФ, Анализ гистограммы восстановленного сигнала, показывает наличие искажения в восстановленном сигнале, которое, однако, визуально никак не определяется. При значении $\delta=15\%$ на данном комплекте растровых изображений был получен коэффициент сжатия – 96%, поскольку с помощью 16 ортов СХПБФ удалось закодировать 360 растровых изображений каналов. Гистограмма распределения яркостей ошибки восстановления приведена на рисунке 3е.

6. Выводы

В работе рассмотрена проблема сокращения информационной избыточности растровых ГСИ ДЗЗ. Отмечено, что использование для этих целей традиционных СБФ (тригонометрические, экспоненциальные системы, системы функций Лагера и пр.) не всегда оправдано, поскольку тогда выбор информативной СБФ ведется чаще всего исходя из внешнего сходства исходного сигнала и СБФ, что может приводить к тому, что оптимальная, либо близкая к ней, СБФ может быть не найдена, особенно в случае сложной формы первоначального сигнала. Разработаны объективные методы выбора СБФ, опирающиеся лишь на свойства совокупности исходных данных. К числу таких методов относится кодирование исходной информации по СХПБФ. В работе приведено описание адаптированного к обработке ГСИ алгоритма формирования СХПБФ, и показана его достаточно высокая эффективность (до 90–96%) устранения информационной избыточности на растровых ГСИ при низком уровне вносимых в сигнал искажений на этапе декодирования.

Авторы планируют продолжить исследовать возможность дальнейшего сокращения объема вычислений при кодировании ГСИ по СХПБФ, использование СХПБФ в задачах подавления помех на ГСИ, а также исследовать возможность построения адаптивных алгоритмов сжатия ГСИ на базе СХПБФ.

7. Литература

- [1] Chang, Ch.-I. Hyperspectral Imaging: Techniques for Spectral Detection and Classification / Ch.-I. Chang // Plenum Publishers. – N.Y.: Kluwer Academic. – 2003. – 370 p.
- [2] Попов, М.А. Методы оптимизации числа спектральных каналов в задачах обработки и анализа данных дистанционного зондирования Земли / М.А. Попов, С.А. Станкевич // Современные проблемы дистанционного зондирования Земли из космоса. – 2006. – Т. 3, № 1. – С. 106-112.
- [3] Lukin, V. Processing of Multichannel RS data for Environment Monitoring, Proceedings of NATO Advanced Research Workshop on Geographical Information / V. Lukin // Processing and Visual Analytics for Environmental Security. – Trento, Italy, Springer Netherlands, 2009. – P. 129-138.
- [4] Kaarna, A. Compression of Spectral Images / A. Kaarna // Vision Systems: Segmentation and Pattern Recognition. – Vienna: I-Tech, 2007. – P. 269-298.
- [5] Yu, G. Image compression systems on board satellites / G.Yu, T. Vladimirova, M.N. Sweeting // Acta Astronautica. – 2009. – Vol. 64. – P. 988-1005.
- [6] Ponomarenko, N.N. Automatic Approaches to OnLand/OnBoard Filtering and Lossy Compression of AVIRIS Images / N.N. Ponomarenko, V.V. Lukin, M.S. Zriakhov, A. Kaarna, J. Astola // Proceedings of IGARSS. – 2008. – Vol. III. – P. 254-257.
- [7] Motta, G. Compression of hyperspectral imagery / G. Motta, F. Rizzo, J.A. Storer // Proceedings of Data Compression Conference. – 2003. – P. 333-342.
- [8] Перспективные информационные технологии дистанционного зондирования Земли: моногр. / под ред. В.А. Сойфера. – Самара: Новая техника, 2015. – 256 с.
- [9] Неймарк, Ю.И. Распознавание образов и медицинская диагностика / Ю.И. Неймарк, З.С. Баталова, Ю.Г. Васин, М.Д. Брейдо. – М.: Наука, 1972. – 328 с.
- [10] Неймарк, Ю.И. Кодирование больших массивов информации в связи с задачами распознавания образов / Ю.И. Неймарк, Ю.Г. Васин // Изв. Высш. уч. зав. – Радиофизика. – 1968. – № 7. – С. 1081-1086.

- [11] Андерсон, Т. Введение в многомерный статистический анализ / Т. Андерсон. – М.: Физматгиз, 1963. – 500 с.
- [12] Christophe, E. Quality criteria benchmark for hyperspectral imagery / E. Christophe, D. Leger, C. Mailhes // IEEE Transactions on Geoscience and Remote Sensing. – 2005. – Vol. 43(9). – P. 2103-2114.
- [13] Christophe, E. Hyperspectral Image Compression: Adapting SPIHT and EZW to Anisotropic 3-D Wavelet Coding / E. Christophe, C. Mailhes, P. Duhamel // IEEE Transactions on Image Processing. – 2008. – Vol. 17(12). – P. 2334-2346.
- [14] Ryan, M.J. A Suitable Distortion Measure for the Lossy Compression of Hyperspectral Data / M.J. Ryan, J.F. Arnold // Proceedings of IGARSS. – 1998. – P. 2056-2058.
- [15] Ponomarenko, N. Preliminary automatic analysis of characteristics of hyperspectral AVIRIS images / N. Ponomarenko, V. Lukin, M. Zriakhov, A. Kaarna // Proceedings of MMET. – Kharkov, Ukraine, 2006. – P. 158-160.
- [16] Pratt, W.K. Digital Image Processing. Fourth Edition / W.K. Pratt. – NY, USA, Wiley-Interscience, 2007. – 807 p.
- [17] Ту, Дж. Принципы распознавания образов / Дж. Ту, Р. Гонсалес. – М.: Мир, 1978. – 411 с.
- [18] Borusyak, A.V. Context - based method for lossless compression of RGB and multispectral images / A.V. Borusyak, P.A. Pakhomov, D.Yu. Vasin, V.E. Turlapov // CEUR Workshop Proceedings. – 2018. – Vol. 2210. – P. 323-329.

Благодарности

Работа выполнена при финансовой поддержке гранта РФФИ № 18-07-00715\18.

Elimination of information redundancy of hyper spectral raster images by method "well adapted" basis

D.Yu. Vasin¹, V.P. Gromov¹, P.A. Pakhomov¹

¹Lobachevsky State University (National Research University), Gagarin Avenue 23, Nizhni Novgorod, Russia, 603952

Abstract. The work is devoted to the consideration of the issues of eliminating information redundancy of hyperspectral raster data (HRD). The characteristic properties of the HRD are listed, a brief description of the existing HRD compression methods is given. The possibility of using local, homogeneous "well-adapted" basic functions (LHWABF) to eliminate information redundancy and adaptive compression of the HRD is considered. An algorithm for constructing a LHWABF system for the HRD based on the Chebyshev approximation is proposed. The results of computational experiments, including the use of a graphics processor, are presented. The effectiveness of the proposed method of adaptive compression HRD is shown.