

# Исследование корреляции эмпирических мод и низкочастотных остатков сигнатур гиперспектральных изображений

В.Е. Турлапов<sup>1</sup>, П.А. Пахомов<sup>1</sup>

<sup>1</sup>Национальный исследовательский государственный университет им. Н.И. Лобачевского, Гагарина 23, Нижний Новгород, Россия, 603952

**Аннотация.** При работе с гиперспектральными изображениями (ГСИ) ДЗЗ приходится иметь дело с шумовыми составляющими, внесенными прежде всего атмосферой. Шум в значительной мере влияет на дальнейшую обработку ГСИ, в том числе на успех определения материала подстилающих поверхностей. В работе исследованы на содержание шума компоненты разложения сигнатуры пикселя ГСИ методом эмпирических мод: эмпирические моды и низкочастотные остатки сигнатур ГСИ, и дальнейшая очистка ГСИ от шума. В качестве исходных данных используется ГСИ Moffett Field спектрометра AVIRIS. Проведена оценка корреляции получаемых эмпирических мод между собой. Исследованы сигнатуры нескольких типов поверхности и их эмпирические моды разных порядков.

## 1. Введение

Настоящая работа является продолжением исследования гиперспектральных изображений и описывает попытки детектирования и фильтрации шумов различной природы, преимущественно шумов, добавленных атмосферой во время фиксации ГСИ. Проблема наличия шумовых компонент часто возникает при работе с ГСИ [3], а сам инструмент - гиперспектральные изображения, становится доступнее и получает широкое распространение в различных областях деятельности человека [1, 4, 5, 8].

В момент фиксации ГСИ на снимок попадает не только информация о яркости определённой точки на поверхности, но и информация о мгновенном состоянии атмосферы на пути луча. Чтобы свести к минимуму влияние состояния атмосферы необходимо знать о значениях таких параметров, как количество водяного пара, распределения аэрозолей и видимость сцены [7]. В большинстве случаев, такая информация недоступна. Поэтому задача получения коэффициентов коррекции шума из сигнатур ГСИ достаточно актуальна.

## 2. Материалы и методы

Метод эмпирических мод хорошо подходит для решения задач анализа данных, в том числе с целью определения наличия посторонних составляющих, например, шумовых. Основная идея метода заключается в получении разных частотных составляющих из исходного сигнала [2]. В результате работы алгоритма построения эмпирических мод из исходного сигнала выделяется набор эмпирических мод разных порядков и низкочастотные остатки. В работе [6] показано, что природа эмпирических мод первых порядков шумоподобна.

Исследование выполнено для всех основных типов объектов, присутствующих на ГСИ Moffett Field. Это ГСИ выбрано по причине присутствия наибольшего, из доступных свободно ГСИ, разнообразия и неоднородности объектов. И, прежде всего, - разнообразия водных объектов. Это территория первого в USA заповедника водно-болотных угодий, который, в силу обстоятельств, расположен внутри городских территорий. Таким образом, это ГСИ наиболее интересно для различения малых вариаций состояния объектов, которое наиболее зависимо от наличия шума.

На рисунке 1 показано пространственное расположение образцов материалов выбранных для исследования. Эти сигнатуры выбраны на основе полноцветного представления ГСИ. Не имея расшифровки материалов поверхности, присутствующих на исследуемом ГСИ, можно сделать лишь предположение о том, какой это материал. Что, собственно, и было сделано. В таблице 3 приведено описание выбранных сигнатур.



Рисунок 1. Сигнатуры материалов для исследования.

### 3. Экспериментальные исследования

В работе [6] показано, что большая часть сигнатур имеет высокий коэффициент корреляции, т.е. общая форма сигнатур достаточно схожа между собой по всему ГСИ. Отличия в них присутствуют лишь в некоторых диапазонах каналов, которые, в свою очередь, являются отличительными признаками материалов подстилающей поверхности. Если в построенной эмпирической моде находится шумовая составляющая, то корреляция между сигнатурой и эмпирической модой должна быть слабой. Однако, корреляция между модами разных порядков должна увеличиваться от каждого предыдущего порядка к следующему.

Для последующих этапов исследования необходимо решить задачу удаления континуума (выпрямления огибающей, определяемой верхними слоями атмосферы). Это позволит убрать общую коррелированность сигнатур, существующую за счет модуляции светового потока верхними слоями атмосферы. В целом, задача сводится к применению некоторого численного фильтра к значениям пикселей.

Необходимо оценить статистические характеристики по каждому каналу ГСИ. Это важно сделать, потому что размерность типа данных пикселя исследуемого в настоящей работе ГСИ составляет 2 Байта. В некоторых точках значение пикселя - его яркость, может достигать существенно больших величин, чем среднее по каналу. Данное явление указывает на наличие очень ярких точек на изображении конкретного канала. К сожалению, нельзя точно сказать, являются ли такие точки бликами отражающих поверхностей или характерными признаками поверхностных материалов, которые, в свою очередь критически важны для дальнейших процедур классификации и распознавания.

Проведём анализ канала 97, как содержащего самый яркий пиксел. Выберем все пиксели большие среднего арифметического по каналу. Всего таких пикселей 6310. Из них для дальнейшего рассмотрения возьмем только те, значения которых больше 4000. Построим среднее по окрестности пикселя (8 соседних с данным пикселом). При этом, значения 0 не

учитываем при подсчёте среднего. Вычислим разницу между значением пиксела и средним арифметическим по его окрестности. Сформируем отношение вычисленной разницы к среднему по окрестности и рассмотрим только те пикселы, для которых полученное отношение больше 1. В данном случае таких пикселов всего 23.

**Таблица 1.** Анализ канала 97.

Значение пиксела	Среднее окрестности	Разница	Отношение
23885	3774	20110	5,33
16053	3239	12814	3,96
10663	4614	6049	1,31
7658	2854	4804	1,68
7031	3338	3693	1,11
7031	2763	4268	1,54
6386	2487	3899	1,57
6339	2536	3802	1,50
6257	1673	4584	2,74
5655	1741	3914	2,25
5515	2029	3485	1,72
5436	1704	3731	2,19
5350	1979	3371	1,70
5135	1785	3350	1,88
5006	1905	3101	1,63
4490	1658	2832	1,71
4324	2106	2217	1,05
4282	1607	2674	1,66
4272	1887	2384	1,26
4161	1267	2894	2,28
4147	1900	2247	1,18
4145	1691	2454	1,45
4037	1647	2389	1,45

Применим такой фильтр ко всему ГСИ. Статистические характеристики по некоторым каналам, в том числе по каналу с максимальным значением в пределах всего ГСИ, до применения фильтра и после приведены в таблице 2. Так как все отрицательные значения в ходе проведения вычислительных экспериментов заменялись значением 0, столбец для минимальных значение не показан.

Из таблицы 2 видно, что применение фильтра позволило значительно понизить общие максимумы по ГСИ. Однако, статистические показатели не изменились значительно, что говорит о сохранении информационной составляющей.

Сигнатуры разных материалов для данного ГСИ раскладываются на разное количество эмпирических мод, от 17 до 21. Чтобы сохранить наглядность материала далее информация будет приводиться только по части эмпирических мод. Моды остальных порядков опущены. Тенденция в них сохраняется.

На рисунке 2 показаны изображения получаемых эмпирических мод и исходная сигнатура. Видно, что форма эмпирических мод представляет собой высокочастотный сигнал с нулевым средним.

Для определения степени коррелированности, как и в работе [6], использован линейных коэффициент корреляции Пирсона. Чтобы отделить эмпирические моды, содержащие шум некоррелированный с полезным сигналом, от мод, несущих составляющие коррелированные с сигналом, исследовалась корреляция эмпирических мод с исходной сигнатурой. В таблице 4 показаны значения коэффициента корреляции между сигнатурой и её модами разных порядков.

**Таблица 2.** Статистические характеристики.

Номер канала	До применения фильтра			После применения фильтра		
	Максимум	Среднее	Отклонение	Максимум	Среднее	Отклонение
1	6788	1208,89	284,89	4275	1208,88	284,83
2	8413	1341,88	320,21	4760	1341,87	320,11
3	9291	1384,97	336,82	9291	1384,96	336,70
4	12233	1708,46	428,46	11461	1708,43	428,18
5	18356	2421,98	632,35	15912	2421,92	631,81
6	20226	2559,77	703,69	17312	2559,68	702,96
7	19714	2428,81	697,35	17168	2428,67	696,37
			...			
95	20520	651,45	442,57	7412	651,38	441,74
96	20021	555,37	376,80	6541	555,32	376,01
97	23885	651,00	440,04	7456	650,93	439,04
98	20223	551,95	372,23	6322	551,90	371,42
99	21258	574,67	388,49	6588	574,62	387,60
			...			
220	7704	94,99	68,96	3869	94,98	68,68
221	7708	81,91	59,31	3470	81,91	58,98
222	5931	53,13	38,52	2325	53,12	38,21
223	2105	21,45	15,68	2105	21,45	15,68
224	2521	24,87	18,22	2521	24,87	18,22

**Таблица 3.** Описание материалов для исследования.

№	Описание материала	Координаты сигнатуры (x,y)
1	Растительность	442; 1332
2	Бетон	447; 1275
3	Грунт	426; 1351
4	Вода чистая	209; 470
5	Вода мутная	315; 775
6	Соляные разработки	540; 630

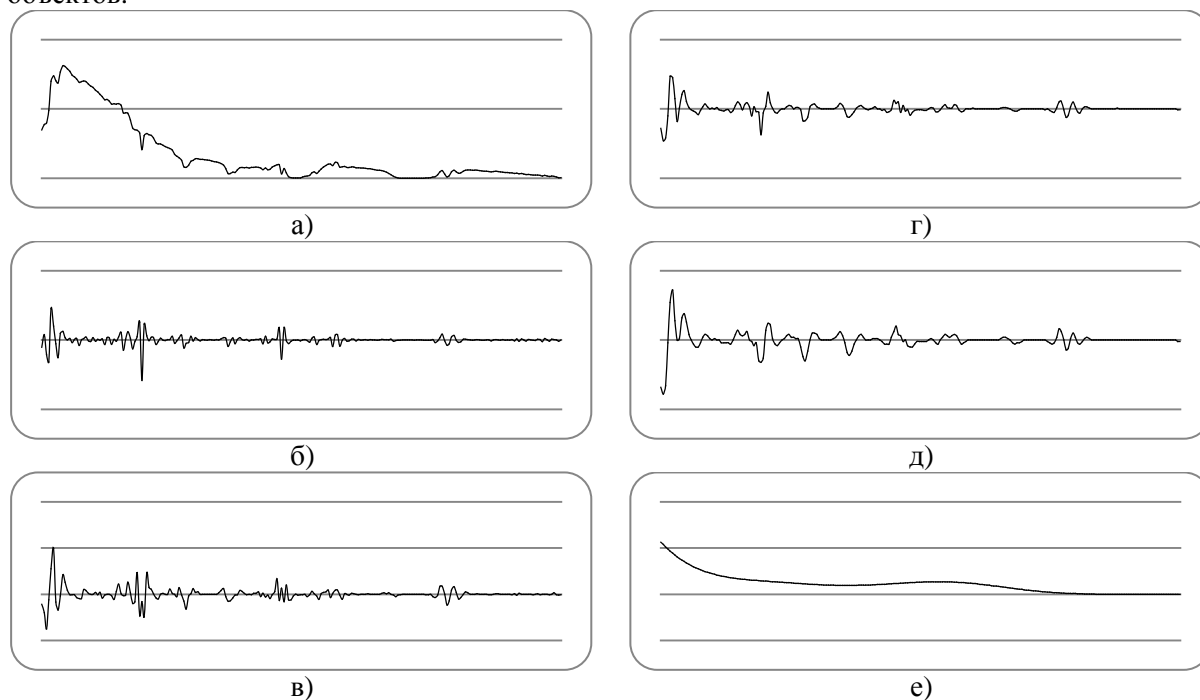
**Таблица 4.** Коррелированность сигнатуры с эмпирическими модами разных порядков.

		Порядок моды					
		1	2	3	4	10	17
Материал	1	0.151	0.199	0.298	0.322	0.494	0.716
	2	0.142	0.166	0.246	0.261	0.394	0.625
	3	0.125	0.162	0.231	0.253	0.457	0.719
	4	0.094	0.135	0.181	0.200	0.336	0.724
	5	0.091	0.126	0.178	0.195	0.357	0.769
	6	0.075	0.105	0.181	0.210	0.487	0.832

Анализируя таблицу 4 можно сделать следующее заключение. Коэффициент корреляции Пирсона для всех мод первого порядка находится в диапазоне 0.075-0.151; второго порядка - в диапазоне 0.105-0.199; 3-го порядка - в диапазоне 0.178-0.298. Величина 0.3 превышена для эмпирической моды 4 уровня только для растительности. Для остальных типов объектов только для мод 6-9 уровней. Согласно шкале Чеддока значения ниже 0.3 соответствуют очень слабой корреляции. Таким образом: моды 1-3 порядков несут в основном шумовую составляющую.

Для выяснения природы шума и подтверждения связанности шума с влиянием атмосферы, исследовалась корреляция между модами одного уровня, но принадлежащими сигнатурам

разных материалов. В таблице 5 показаны матрицы взаимной корреляции мод для разных объектов.



**Рисунок 2.** Разложение сигнатуры на эмпирические моды. а - сигнатура. б-д - эмпирические моды 1-4 соответственно. е - последняя эмпирическая мода.

**Таблица 5.** Коррелированность мод разных материалов.

		Материал						Материал					
		5	6	4	3	2	1	5	6	4	3	2	1
		<b>Мода порядка 1</b>						<b>Мода порядка 7</b>					
Материал	5	1,00	0,87	0,96	0,91	0,73	0,78	1,00	0,76	0,98	0,87	0,66	0,35
	6	0,87	1,00	0,80	0,89	0,81	0,88	0,76	1,00	0,70	0,67	0,52	0,09
	4	0,96	0,80	1,00	0,85	0,68	0,71	0,98	0,70	1,00	0,83	0,60	0,32
	3	0,91	0,89	0,85	1,00	0,89	0,89	0,87	0,67	0,83	1,00	0,90	0,60
	2	0,73	0,81	0,68	0,89	1,00	0,87	0,66	0,52	0,60	0,90	1,00	0,75
	1	0,78	0,88	0,71	0,89	0,87	1,00	0,35	0,09	0,32	0,60	0,75	1,00
		<b>Мода порядка 2</b>						<b>Мода порядка 9</b>					
Материал	5	1,00	0,85	0,98	0,88	0,70	0,68	1,00	0,67	0,97	0,85	0,66	0,28
	6	0,85	1,00	0,79	0,81	0,73	0,76	0,67	1,00	0,61	0,65	0,50	0,03
	4	0,98	0,79	1,00	0,84	0,65	0,61	0,97	0,61	1,00	0,77	0,60	0,26
	3	0,88	0,81	0,84	1,00	0,89	0,81	0,85	0,65	0,77	1,00	0,87	0,50
	2	0,70	0,73	0,65	0,89	1,00	0,83	0,66	0,50	0,60	0,87	1,00	0,71
	1	0,68	0,76	0,61	0,81	0,83	1,00	0,28	0,03	0,26	0,50	0,71	1,00
		<b>Мода порядка 3</b>											
Материал	5	1,00	0,84	0,99	0,87	0,68	0,57						
	6	0,84	1,00	0,78	0,75	0,65	0,56						
	4	0,99	0,78	1,00	0,83	0,62	0,52						
	3	0,87	0,75	0,83	1,00	0,90	0,75						
	2	0,68	0,65	0,62	0,90	1,00	0,81						
	1	0,57	0,56	0,52	0,75	0,81	1,00						

Из таблицы 5 видно, что матрица корреляции для первой моды содержит величины коэффициентов корреляции первой моды разных объектов от 0.68 до 0.96 при среднем 0.86. С

возрастанием номера моды величины корреляции падают и для второй и третьей моды их средние величины составляют соответственно 0.82 и 0.78, оставаясь при этом достаточно высокими. Далее идет плавный спад средних: так, 7-я мода - 0.7; 9-я мода - 0.66; и т.д.

#### 4. Заключение

Проведена оценка корреляции эмпирических мод между собой. Исследована корреляция сигнатуры со своими эмпирическими модами разных порядков. Исследованы сигнатуры нескольких поверхностных материалов. Приведены величины корреляции. Установлено, что коррелированность первых 3-х эмпирических мод разных материалов достаточно высока. Это значит, что в эмпирических модах первых 3-порядков содержится похожая информация, даже если эти моды были получены путём разложения сигнатур разных материалов. Указанное свидетельствует в пользу того, что шумовая составляющая сигнатур разных объектов, заключенная в первых двух-трех модах, обусловлена, в силу высокой коррелированности, одним источником, а именно, влиянием атмосферы.

#### 5. Благодарности

Работа выполнена при поддержке гранта РФФИ № 16-11-00068.

#### 6. Литература

- [1] Аншаков, Г.П. Использование мульти- и гиперспектральных данных дистанционного зондирования для автоматизированного мониторинга рек и водоёмов в весенний период / Г.П. Аншаков, Ю.Н. Журавель, А.В. Ращупкин // Компьютерная оптика. – 2015. – Т. 39, № 2. – С. 224-233. DOI: 10.18287/0134-2452-2015-39-2-224-233.
- [2] Гурьянов, Ф.А. Оптимизация корреляционных методов регистрации биомедицинских изображений / Ф.А. Гурьянов, А.С. Крылов // Труды 27-й Международной конференции по компьютерной графике и машинному зрению, 2017. – С. 253-257.
- [3] Денисова, А.Ю. Атмосферная коррекция гиперспектральных изображений с использованием малого объёма верифицированных данных / А.Ю. Денисова, В.В. Мясников // Компьютерная оптика. – 2016. – Т. 40, № 4. – С. 526-534. DOI: 10.18287/2412-6179-2016-40-1-526-534.
- [4] Козодеров, В.В. Обработка и интерпретация данных гиперспектральных аэрокосмических измерений для дистанционной диагностики природно-техногенных объектов // Исследование Земли из космоса. – 2009. – № 2. – С. 36-54.
- [5] Лебедев, Л.И. Комплексный анализ и мониторинг состояния окружающей среды на основе данных ДЗЗ / Л.И. Лебедев, Ю.В. Ясаков, Т.Ш. Утешева, В.П. Громов, А.В. Борусяк, В.Е. Турлапов // Компьютерная оптика. – 2019. – Т. 43, № 2. – С. 282-295. DOI: 10.18287/2412-6179-2019-43-2-282-295.
- [6] Пахомов, П.А. Исследование сигнатур пикселей гиперспектрального изображения методом эмпирических мод / П.А. Пахомов, А.В. Борусяк, В.Е. Турлапов // Сборник трудов международной конференции и молодежной школы "Информационные технологии и нанотехнологии" (ИТНТ) - Самара: Новая техника, 2018. – С. 1131-1141.
- [7] Токарева, О.С. Обработка и интерпретация данных дистанционного зондирования Земли: учебное пособие / О.С. Токарева – Томск: Издательство Томского политехнического университета, 2010.
- [8] Чабан, Л.Н. Исследование возможностей классификации растительного покрова по гиперспектральным изображениям в пакетах тематической обработки данных дистанционного зондирования / Л.Н. Чабан, Г.В. Вечерук, Т.С. Гаврилова // Труды Московского физико-технического института. – 2009. – Т. 3, № 1. – С. 171-180.

## Investigation of correlation of empirical modes and low-frequency residues of hyperspectral image signatures

V.E. Turlapov<sup>1</sup>, P.A. Pakhomov<sup>1</sup>

<sup>1</sup>National Lobachevsky State Research University, Gagarin ave. 23, Nizhny Novgorod, Russia, 603952

**Abstract.** When working with hyperspectral images (GSI) of remote sensing data, it is necessary to deal with noise components introduced primarily by the atmosphere. Noise significantly affects the further processing of the GSI, including the success of determining the material of the underlying surfaces. In this paper, the components of the decomposition of the GSI pixel signature by the method of empirical modes are studied for the noise content: empirical modes and low-frequency residues of the GSI signatures, and further purification of the GSI from noise. As the initial data, the Moffett Field GSI of the AVIRIS spectrometer is used. The correlation between the obtained empirical modes is estimated. Signatures of several surface types and their empirical modes of different orders are investigated.