

УДК 004.93

НЕЙРОСЕТЕВАЯ КЛАССИФИКАЦИЯ ГИПЕРСПЕКТРАЛЬНЫХ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ПРЕОБРАЗОВАНИЯ ГИЛЬБЕРТА — ХУАНГА

Е. С. Нежевенко, А. С. Феоктистов, О. Ю. Дашевский

*Институт автоматизации и электрометрии СО РАН,
630090, г. Новосибирск, просп. Академика Коптюга, 1
E-mail: nejevenko@iae.nsk.su*

Метод классификации изображения с предварительной трансформацией его к главным компонентам и использованием преобразования Гильберта — Хуанга исследуется на примере нейросетевой классификации гиперспектрального изображения. Показана эффективность метода путём сравнения с традиционными методами нейросетевой классификации при использовании в качестве признаков спектральных компонент и главных компонент без привлечения пространственной информации. Для классификации применяются нейронные радиально-базисные и комплексные сети.

Ключевые слова: классификация, гиперспектральное изображение, преобразование Гильберта — Хуанга, главные компоненты, комплексные нейронные сети.

DOI: 10.15372/AUT20170209

Введение. Гиперспектральная съёмка — перспективный метод дистанционного зондирования Земли. Каждому пикселу гиперспектрального изображения (ГСИ) соответствует вектор, компоненты которого есть значения отражения или излучения света в очень узком участке спектра. В результате всё изображение представляет собой гиперкуб, сечения которого являются изображениями в узком спектральном диапазоне. Анализируя пиксели, необходимо определить, к какому классу относится тот или иной участок изображения. Для этого формируются так называемые зоны интереса, т. е. участки, где тем или иным способом (наиболее предпочтительно маршрутной съёмкой) установлен класс. Перечень классов определяется характером ландшафта (любой лес, луг, вода и т. д.). Затем каждая зона интереса разбивается на две выборки: обучающую и тестовую. С помощью обучающей выборки настраивается алгоритм классификации, а с помощью тестовой проверяется его эффективность. После настройки (обучения) и проверки алгоритм используется для классификации неизвестных изображений. В качестве первичных признаков можно брать компоненты вектора пиксела ГСИ. Однако эти компоненты сильно коррелированы вследствие того, что спектры отражения или излучения участков поверхности Земли, как правило, являются гладкими функциями, кроме того, значительные участки спектров различных материалов могут быть схожи друг с другом. Отсюда следует, что для классификации такие спектры переизбыточны и необходимо формировать эффективную систему признаков.

Во многих работах, например [1], показано, что наиболее эффективным методом сокращения пространства признаков является метод главных компонент (ГК), т. е. преобразование Кархунена — Лоэва и Хотеллинга. Такое преобразование проводится попиксельно, при этом число признаков, используемых для распознавания, существенно сокращается. Однако, если при попиксельной классификации не учитывается пространственная струк-

тура изображения, это может привести к тому, что, например, тайга и мелколесье будут отнесены к одному классу при схожести спектральных диапазонов отражения. В работах [2, 3] показано, что сочетание попиксельных и пространственных признаков существенно повышает вероятность правильного распознавания фрагментов ГСИ. В подавляющем большинстве методов классификации с помощью пространственных признаков вид преобразования назначается априорно, без учёта специфики изображения. Например, усреднение в окрестности пиксела, "Majority analysis" [4], сглаживание данных посредством медианной фильтрации, вейвлет-анализ — все эти методы не зависят от характеристик самого изображения. Исследуемый в [2] подход, основанный на использовании для классификации каждого пиксела его окрестности, повышает точность классификации на 15–17 %, однако логично предположить, что вид выбираемой окрестности будет зависеть от характера распознаваемой текстуры.

В данной работе предлагается и исследуется эффективность нового (для задачи классификации ГСИ) метода нелинейной трансформации исходного пространства признаков (векторов спектральных компонент ГСИ) на основе преобразования Гильберта — Хуанга (ННТ — Hilbert-Huang Transformation), причём он применяется не к исходному изображению, а к его главным компонентам. Этот метод характерен тем, что может работать с нелинейными и нестационарными функциями [5], какими описываются в большинстве случаев ГСИ поверхности Земли.

Преобразование Гильберта — Хуанга. Поскольку преобразование Гильберта — Хуанга не нашло ещё широкого применения при обработке изображений, дадим описание его основ. Для простоты представим одномерные функции. Преобразование состоит из двух этапов: разложение на эмпирические моды (преобразование Хуанга) и спектральный анализ Гильберта. Второй этап позволяет найти мгновенные амплитуды и частоты сигнала, которые его характеризуют. В основе преобразования Хуанга лежит предположение, что любые данные состоят из разнообразных видов колебаний, наложенных друг на друга. Результат — конечное сложное колебание. На первый взгляд трансформация аналогична преобразованию Фурье, однако она существенно отличается как от этого преобразования, так и от вейвлет-преобразования, близкого к преобразованию Хуанга: при ННТ используются внутренние колебания, присущие анализируемому сигналу, а не фиксированный набор частот или вейвлет-функций. Функции базиса разложения (IMF — Intrinsic Mode Function, эмпирические моды) имеют следующие особенности:

- 1) в эмпирической моде число экстремумов и число нулевых пересечений должно быть одинаковым или отличаться, самое большее, на единицу;
- 2) в любой точке эмпирической моды среднее значение огибающих, определённых локальными максимумами и минимумами, равно нулю.

Алгоритм разложения сигнала на эмпирические моды состоит из нескольких этапов.

1. Определяются локальные экстремумы сигнала (минимумы и максимумы).
2. Получаются огибающие сигнала путём аппроксимации отдельно локальных максимумов и отдельно локальных минимумов кривой аппроксимации. Возможны различные методы аппроксимации. Так, Хуанг предлагал для этой процедуры использовать кубические сплайны, однако для двумерных изображений такая интерполяция даёт сильные возмущения на краях изображений. Поэтому в данном случае предпочтительно применение для интерполяции радиально-базисных функций (RBF — Radial Basis Function).

3. Вычисляются средние значения $m_1(t)$ верхней и нижней огибающих и разность сигнала и среднего значения: $h_1(t) = x(t) - m_1(t)$.

4. Далее в качестве исходного сигнала берётся уже функция $h_1(t)$ и алгоритм повторяется до выполнения условия останова. Используются два различных критерия останова.

Первый определён как нормализованная квадратичная разность двух последовательных операций отсеивания:

$$SD_k = \sum_{t=0}^T |h_{k-1}(t) - h_k(t)|^2 / \sum_{t=0}^T h_{k-1}^2(t).$$

Если квадратичная разность SD_k меньше, чем предопределённое значение, процесс отсеивания будет остановлен. Второй критерий (S) заключается в том, что число экстремумов на последующей и предыдущей итерациях должно остаться неизменным.

5. После останова $h_k(t)$ принимается за первую IMF-функцию $c_1(t)$. В качестве исходного сигнала берётся разность $r_1(t) = x(t) - c_1(t)$, и процесс повторяется n раз до тех пор, пока остаток $r_n(t)$ не станет либо константой, либо монотонной функцией.

6. В конечном итоге мы имеем разложение сигнала на эмпирические моды в виде

$$x(t) = \sum_{j=1}^n c_j(t) + r_n(t).$$

Завершающим шагом ННТ является преобразование Гильберта, возможность которого основана на том, что эмпирические моды, полученные в ходе разложения, являются аналитическими сигналами. Это преобразование позволяет выделить такие признаки, характеризующие сигнал, как мгновенные частоты и амплитуды каждой моды.

Для действительного сигнала $s(t)$ преобразование Гильберта определяется главным значением (PV) интеграла

$$H_s(t) = PV \int_{-\infty}^{\infty} \frac{s(\tau)}{t - \tau} d\tau, \quad Z(t) = s(t) + iH_s(t) = a(t)e^{i\theta(t)},$$

$$a(t) = \sqrt{s(t)^2 + H_s(t)^2}, \quad \theta(t) = \arctan \left[\frac{H_s(t)}{s(t)} \right],$$

где $a(t)$ и $\theta(t)$ — амплитуда и фаза сигнала. Мгновенная частота может быть вычислена по формуле

$$\omega(t) = d\theta(t)/dt.$$

Соответственно при применении преобразования Гильберта — Хуанга к изображению для каждого пиксела возникает $2m$ признаков (m — число полученных эмпирических мод [5]). Пример разложения на эмпирические моды двумерного изображения (не гиперспектрального) представлен на рис. 1: исходное изображение (первый фрагмент) и три его эмпирические моды.

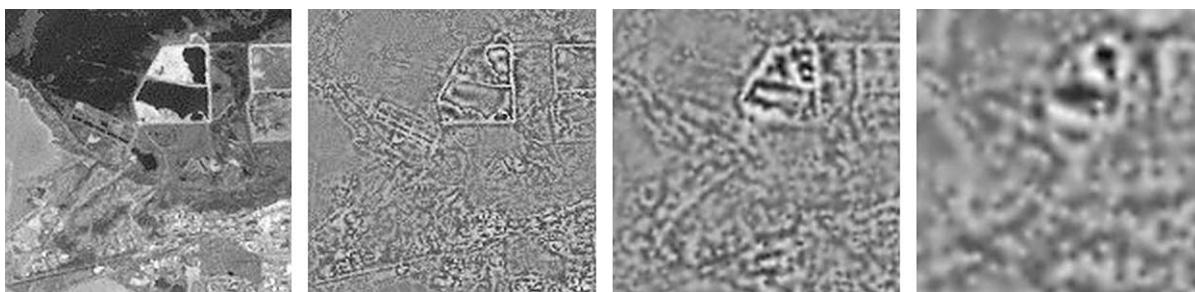


Рис. 1. Фрагмент классифицируемого изображения полигона «Индиан Пайнс» (штат Индиана, США) и три его эмпирические моды

Очевидно, что применять ННТ к полному ГСИ не следует: достаточно большое число признаков каждого пиксела, равное числу спектральных компонент (а для типичного ГСИ это сотни), увеличивается в несколько раз. Переход к ГК устанавливает иерархию признаков: чем выше номер главной компоненты, тем менее она значима при распознавании. Что касается значимости номера эмпирической моды для распознавания — данная проблема требует дополнительного исследования. Поэтому наша задача — выяснить, какая минимальная комбинация главных компонент и эмпирических мод даёт достаточно высокую вероятность распознавания при классификации ГСИ.

Эксперимент. Впервые метод классификации ГСИ с помощью преобразования Гильберта — Хуанга описан в [6]. Однако эксперимент проводился на мультиспектральном изображении, где число признаков каждого пиксела измеряется десятками, кроме того, обучающая выборка формировалась путём визуального (экспертного) анализа изображения. В данной работе для экспериментального исследования взяты широко известные тестовые фрагменты ГСИ, предлагаемые в открытом пакете MultiSpec [7] (эти же изображения использовались в [2, 8]). Изображение получено в рамках программы AVIRIS (Airborne Visible Infrared Imaging Spectrometer) на опытном полигоне «Индиан Пайнс». Фрагмент имеет размер 145×145 пикселей, разрешение 20 м/пиксел, число спектральных компонент 220 в диапазоне 0,4–2,5 мкм, при этом 20 компонент, обладающих высоким уровнем шума, при проведении исследований удалены из рассмотрения. Обучающая выборка формировалась путём маршрутной съёмки. Изображение и его маршрутная классификация представлены на рис. 2.

Данное изображение было преобразовано к ГК, после чего по критерию «каменистая осыпь» (графику убывания собственных значений) для классификации отбиралось пять главных компонент, обеспечивающих 97,32 % дисперсии. К каждой из ГК применено ННТ и получено по пять эмпирических мод, мгновенные амплитуды которых взяты в качестве признаков. Результат преобразований показан на рис. 3: каждая строка — это главная компонента (первый фрагмент) и её пять эмпирических мод.

Таким образом, каждый пиксел изображения характеризовался 25 признаками. Преобразование Гильберта — Хуанга проведено в программе, разработанной в среде MATLAB, основными модулями которой являются модули ННТ; для поиска локальных экстремумов реализованы операторы морфологической реконструкции, а для построения огибающих — аппроксимация с помощью RBF. Для определения останова использовались критерии S и SD_{\max} .

Классификация проводилась в двух типах нейронных сетей: в бинарных сетях стандартной программы StatSoft и в комплекснозначной сети с многоуровневыми нейронами.



Рис. 2. Результаты маршрутной классификации фрагмента

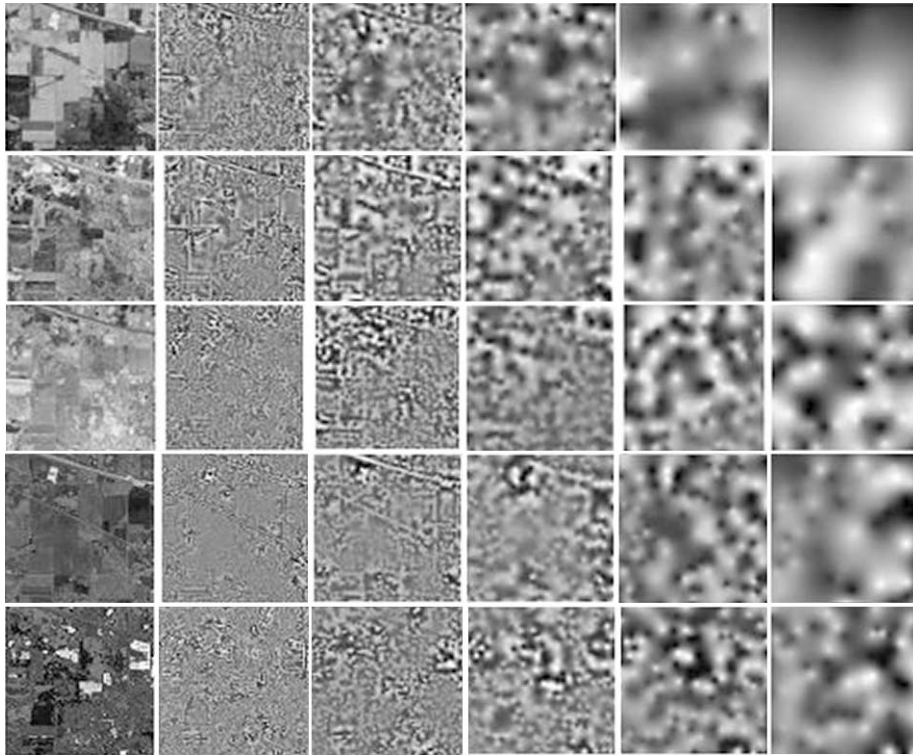


Рис. 3. Результаты преобразования Хуанга

Выбор второго типа обусловлен тем, что в работе [9] на мультиспектральном изображении на обоих типах сетей был получен практически одинаковый результат, но комплекснозначная нейронная сеть имеет некоторые потенциальные преимущества перед бинарной при обучении [10].

При классификации в программе StatSoft конструировалось десять нейронных сетей, из которых отбиралось пять наилучших. Каждый массив делился случайным образом на три выборки: обучающую, контрольную и тестовую. Контрольная выборка использовалась для останова процесса обучения. В результате было получено три массива, подлежащих классификации: массив спектральных компонент (МСК), массив главных компонент (МГК) и массив эмпирических мод главных компонент (МЭМГК). Решалась задача определения вероятности правильной классификации на каждом из обучающих и тестовых массивов и выбора эффективной и наиболее экономичной системы признаков.

В программе StatSoft обучение проводилось в следующих нейронных сетях: линейная сеть, многослойный персептрон, сети на основе радиально-базисных функций (RBF-сети) с различным количеством элементов. Наилучшие результаты обучения и классификации получены в RBF-сети. Один из скриншотов программы StatSoft представлен на рис. 4.

Вероятности распознавания: $P_{об}$ — производительность обучения, P_T — тестовая производительность. На МСК (RBF-сеть из 121 элемента) $P_{об} = 0,756$, $P_T = 0,752$. Как видно, почти 25 % пикселей классифицированы неправильно. На массиве главных компонент (RBF-сеть из 630 элементов) $P_{об} = 0,785$, $P_T = 0,720$, т. е. классификация по ГК практически не отличается от классификации по исходным признакам, правда, число признаков при этом в десятки раз уменьшается. При использовании пространственной информации, когда в качестве признаков взяты амплитуды первых пяти эмпирических мод ГК, при обучении в RBF-сети из 824 элементов получены $P_{об} = 0,997$, $P_T = 0,991$, т. е. вероятность правильной классификации ~ 100 %. А вот обучение и классификация того же изображения с теми же преобразованиями в комплекснозначной сети дало не такие хорошие резуль-

N	Архитектура	Производительность		Тест.		Ошибка обучения	Контрольная ошибка	Тестовая ошибка	Обучение
		обуч.	производительность	производительность	производительность				
4	РБФ 5-5-549-15-1	0,981268	0,964481	0,970726	0,072818	0,082930	0,081403		
5	РБФ 5-5-824-15-1	0,994732	0,983216	0,987900	0,054783	0,067036	0,064822		
8	РБФ 5-5-365-15-1	0,982439	0,975020	0,976971	0,074027	0,078032	0,077689		
9	РБФ 5-5-549-15-1	0,995902	0,986729	0,989071	0,058958	0,063926	0,061215		
10	РБФ 5-5-824-15-1	0,996683	0,991803	0,991023	0,045327	0,052685	0,053420		

Рис. 4. Скриншот результатов работы нейропрограммы StatSoft по классификации фрагмента гиперспектрального изображения

таты: на МГК (архитектура 5-23-16: 5 входных нейронов, 23 скрытых и 16 выходных) $P_T = 0,623$. Для МЭМГК (пять мод в комплекснозначной нейронной сети, архитектура 5-36-16) $P_T = 0,687$. Очевидно, что обучение в комплекснозначной сети не смогло сформировать тех сложных разделяющих гиперплоскостей, которые сформированы RBF-сетью бинарных нейронов из 824 элементов.

Заключение. Таким образом, данной работой мы опровергли наше же предположение, высказанное в [11], о том, что почти стопроцентная правильная классификация изображений с использованием ННТ возможна тогда, когда области изображения, принадлежащие различным классам, разделены в пространстве зазорами и «расплывание» мод не влияет на классификацию, поскольку из рис. 2 видно, что области различных классов почти вплотную примыкают друг к другу. Отсюда следует, что преобразование Гильберта — Хуанга в сочетании с главными компонентами — весьма эффективное преобразование ГСИ и при обучении в нейронной сети обеспечивает очень высокую вероятность правильной классификации. Однако здесь ещё требуются серьёзные исследования при определении эффективного состава мод для разных типов изображений и зон интереса.

СПИСОК ЛИТЕРАТУРЫ

1. **Козик В. И., Нежевенко Е. С.** Выбор эффективной системы признаков при сегментации гиперспектральных изображений // Матер. Междунар. науч.-практ. конф. «Перспективы развития науки и образования». Тамбов: Изд-во ООО «Консалтинговая компания Юком», 2014. Ч. 7. С. 70–72.
2. **Борзов С. М., Потатуркин А. О., Потатуркин О. И., Федотов А. М.** Исследование эффективности классификации гиперспектральных спутниковых изображений природных и антропогенных территорий // Автометрия. 2016. **52**, № 1. С. 3–14.
3. **Plaza A., Benediktsson J. A., Boardman J. W. et al.** Recent advances in techniques for hyperspectral image processing // Remote Sensing of Environment. 2009. **113**, Suppl. 1. P. S110–S122.
4. **Lillesand T. M., Kiefer R. W., Chipman J. W.** Remote Sensing and Image Interpretation. N. Y.: John Wiley & Sons, 2004. 763 p.
5. **Huang N. E., Shen Z., Long S. R. et al.** The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis // Proc. Royal Soc. London. Ser. A. 1998. **454**, Is. 1971. P. 903–995.
6. **Феоктистов А. С., Нежевенко Е. С.** Классификация гиперспектральных изображений с помощью преобразования Гильберта — Хуанга // Интерэкспо Гео-Сибирь. 2015. **4**, № 2. С. 23–27.
7. **MultiSpec.** A Freeware Multispectral Image Data Analysis System. URL: <https://engineering.purdue.edu/~biehl/MultiSpec/hyperspectral.html> (дата обращения: 08.02.2017).

8. **Фурсов В. А., Бибииков С. А., Байда О. А.** Тематическая классификация гиперспектральных изображений по показателю сопряжённости // Компьютерная оптика. 2014. **38**, № 1. С. 154–158.
9. **Дашевский О. Ю., Нежевенко Е. С.** Классификация гиперспектральных изображений с помощью нейронных сетей с бинарными и многоуровневыми нейронами // Интерэкспо Гео-Сибирь. 2015. **4**, № 2. С. 62–66.
10. **Aizenberg I.** MLMVN with soft margins learning // IEEE Trans. Neural Networks and Learning Systems. 2014. **25**, Is. 9. P. 1632–1644.
11. **Феоктистов А. С., Нежевенко Е. С.** Исследование эффективности нейросетевой классификации гиперспектральных изображений с использованием преобразования Гильберта — Хуанга // Интерэкспо Гео-Сибирь. 2016. **4**, № 1. С. 59–62.

Поступила в редакцию 21 декабря 2016 г.
