

**ИССЛЕДОВАНИЕ ПРАКТИЧЕСКОГО ПРИМЕНЕНИЯ МЕТОДА БЫСТРОГО
АНАЛИЗА НЕЗАВИСИМЫХ КОМПОНЕНТ В КЛАССИФИКАЦИИ
МУЛЬТИСПЕКТРАЛЬНЫХ ДАННЫХ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ
ЗЕМЛИ**

Каримова Гульмира Токтомуратовна - старший преподаватель кафедры «Информационные системы и технологии в телекоммуникациях», Институт электроники и телекоммуникации при КГТУ им. И. Раззакова, Кыргызская Республика, 720044, г. Бишкек, пр. Ч. Айтматова 66, e-mail: k.gulpeace@gmail.com

Каримов Бактыбек Токтомурастович – к.т.н., профессор кафедры «Радиоэлектроники», Институт электроники и телекоммуникации при КГТУ им. И. Раззакова, Кыргызская Республика, 720044, г. Бишкек, пр. Ч. Айтматова 66, e-mail: karimov_bt@mail.ru

Аннотация. Классификация данных дистанционного зондирования Земли с целью выявления изменений в почвенно-растительном покрове являются одной из актуальных задач, который ставит перед собой цифровая обработка данных. В статье исследуется метод быстрого анализа независимых компонент (АНК), который применяется при «слепом» разделении множественного спектрального изображения. Рассматриваются принципы анализа независимых компонент, и предлагается применить алгоритм быстрого – АНК для данных Landsat ETM +. В результате работы алгоритма был сжат большой объем данных и получена тематическая информация, а также извлечены области снимка с изменениями в значениях спектрального излучения.

Ключевые слова: Анализ независимых компонент, данные дистанционного зондирования земли, классификация мультиспектральных снимков

INVESTIGATION OF FAST INDEPENDENT COMPONENT ANALYSIS PRACTICAL APPLICATION IN MULTISPECTRAL REMOTE SENSING DATA CLASSIFICATION

Karimova Gulmira Toktomuratovna – Senior Lecturer, Department of Information System and Technology in Telecommunication, Electronics and Telecommunication Institute under Kyrgyz State Technical University names after I.Razzakov, 66 Prospect Ch.Aitmatova, Bishkek, 720044, Kyrgyz Republic, e-mail: k.gulpeace@gmail.com

Karimov Baktybek Toktomuratovich – Candidate of Technical Science, Professor of “Radio Electronics” Department, Electronics and Telecommunication Institute under the Kyrgyz State Technical University named I.Razzakov, 66, Ch.Aitmatov Prospect, Bishkek, 720044, Kyrgyz Republic. E-mail: karimov_bt@mail.ru

Abstract. The classification of remote sensing data in order to detect land cover changes is one of the urgent tasks that digital data processing sets itself. The article explores the method of Independent Component Analysis (ICA), which used in the "blind" separation of multiple spectral images. The ICA principles considered and proposed to apply the fast-ICA algorithm for Landsat ETM + data. The algorithm helps to compress a large amount of data in order to obtain thematic information and changes in image.

Key words: Independent Component Analysis, remote sensing data, multispectral images classification

Введение

Поиск быстрого и высокоточного классификационного метода для выявления изменений в почвенно-растительном покрове является актуальной задачей для специалистов в области цифровой обработки данных [9]. С ростом категорий классификаций, возрастает и вероятность роста ошибки при классификации этих категорий. Возможность с высокой точностью разделить значение пикселей в спектральных каналах, снизит вероятность ошибок при классификации и увеличит скорость обработки.

До сих пор не найдены унифицированные решения в этой области. Существующие методы требуют предварительной обработки изображений, и ограничены в способности точно выявить изменения. Одним из таких методов является метод главных компонент, который требует нахождения пороговых значений для обнаружения изменений. Метод быстрого

анализа независимых компонент используется для дифференцирования сигнала в области медицины [8]. В данной работе сделана попытка применения АНК при классификации спутникового снимка ЕТМ + имеющего три канала и проверки эффективности и преимущества АНК алгоритма в классификации и выявлении изменений в мультиспектральном изображении.

Анализ независимых компонент (Independent Component Analysis)

Анализ независимых компонент (Independent Component Analysis) является статистическим методом разложения сложного набора данных на независимые подразделы [1,2]. Он базируется на «слепом» разделении данных и преобразует наблюдаемый многомерный вектор в компоненты, которые статистически независимы друг от друга настолько, насколько это возможно. Существует по крайней мере три разных определения АНК.

Определение 1: АНК случайного вектора X состоит в нахождении линейного преобразования $S = WX$, так чтобы компоненты S_i были независимыми настолько насколько возможно, в свете максимизации некоторых функций $F(S_1, \dots, S_m)$, которая измеряет независимость.

Определение 2: АНК случайного вектора X состоит из оценки следующей модели, генерирующей данных:

$$X=AS+n$$

где X – исследуемый вектор сигналов, n – вектор шума, A – оцениваемая смесевая матрица и S - взаимно независимые компоненты.

Определение 3: АНК без помех

$$X=AS$$

где A - матрица и S - взаимно независимые компоненты.

Существует множество алгоритмов для реализации АНК, но самым эффективным на сегодняшний день является алгоритм быстрого АНК, который был предложен А.Хювяринен [8], это алгоритм с фиксированной точкой, основанный на оптимизации энтропийной функции, называемой отрицательной энтропией.

В отличие от анализа главных компонент, АНК можно рассматривать как инструмент, основанный на более высоком порядке статистики, и он не только не коррелирует входные сигналы, но также делает результат как можно более независимым, неповторяющимися друг друга [2]. Алгоритм быстрого АНК обладает свойствами объективной оптимизации функций и стабилен. Модель быстрого анализа независимых компонент представлен на рисунке 1.

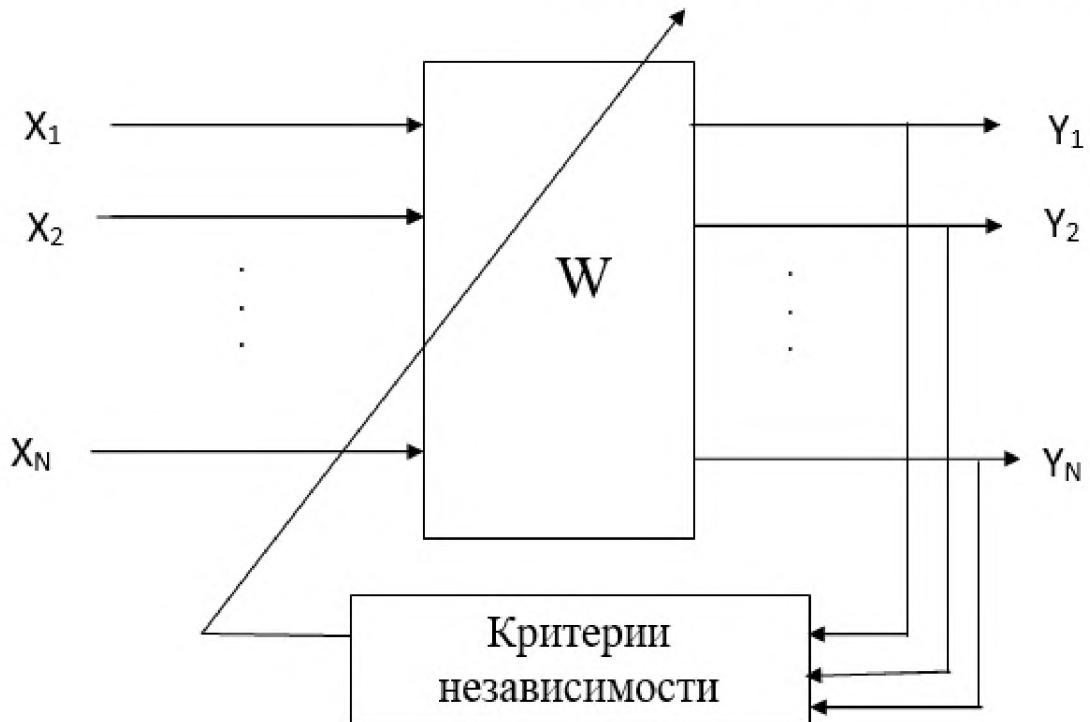


Рис.1 Модель быстрого анализа независимых компонент

$$J_G(w_i) = [E\{G(w_i^T x)\} - E\{G(v)\}]^2, \quad (1)$$

где информативность признака с номером i переменной $y_i = w_i^T x$ и w – m - мерный вектор ограниченный функцией $E\{G(w_i^T x)\} = 1$, является весовым компонентом, v – Гаусовская нулевая случайная величина и единичная дисперсия, $G(y_i)$ – пространство признаков. Для того чтобы достичь идеального разделения канала, нам необходимо следующее условие ограничения

$$E\{(w_k^T x)(w_l^T x)\} = \delta_{jk} \text{ при максимизации } \sum_{i=1}^N J_G(w_i)$$

Изображения полученные в результате поворота в m -мерном пространстве имеют наибольший контраст, они содержат наибольшее количество информации, которое ранее находились в исходных изображениях и были рассреточены на разных каналах. [6]

Классификация изображений

Если предположить, что состав каждого смесового пикселя является случайным значением, и спектральные кривые различных типов объектов земли не зависят друг от друга, каждая кривая спектрального отклика представляет собой полосу источника, а мультиспектральное изображение можно рассматривать как смесовую полосу, полученную из исходных полос снимка, то классификация спектрального изображения становится задачей «слепого» разделения смесовой полосы (канала), для которой и будет использован алгоритм быстрого АНК [5].

Этапы использования алгоритма быстрого АНК для классификации мультиспектрального изображения заключается в следующем:

Шаг 1. Выравнивание (эквализация) гистограммы;

Шаг 2. Удаление шума в полосе (канале);

Шаг 3. Зададим значение весовому вектору w и установим ошибки сходимости ϵ ;

Шаг 4. Обновление весового вектора w . Для улучшения стабильности добавим размер шага μ и используем следующую итерационную формулу (2):

$$w_i^+ = w_i - \mu[E\{xg(w_i^T x)\} - \beta w_i]/[E\{g(w_i^T x)\} - \beta], \quad (2)$$

где $\beta = E\{w_i^T, xg(w_i^T, x)\}$

Шаг 5. Нормализуем вектор веса w_i^+ ;

Шаг 6. Если $|w_{k+1} - w_k| > \varepsilon$, алгоритм не достигает сходимости, повторим шаги (4) и (5);

Шаг 7. Если алгоритм не достигает сходимости и итерация превышает предварительно заданное наибольшее число (например, 100), уменьшаем размера шага на половину и возвращаемся на шаги (4) и (5) до выполнения условия $|w_{k+1} - w_k| < \varepsilon$,

Шаг 8. Получаем матрицу разделения по информативности w и АНК-преобразованный исходный снимок;

Шаг 9. Извлекаем тематическую информации из АНК снимков.

Получение изменений на снимке

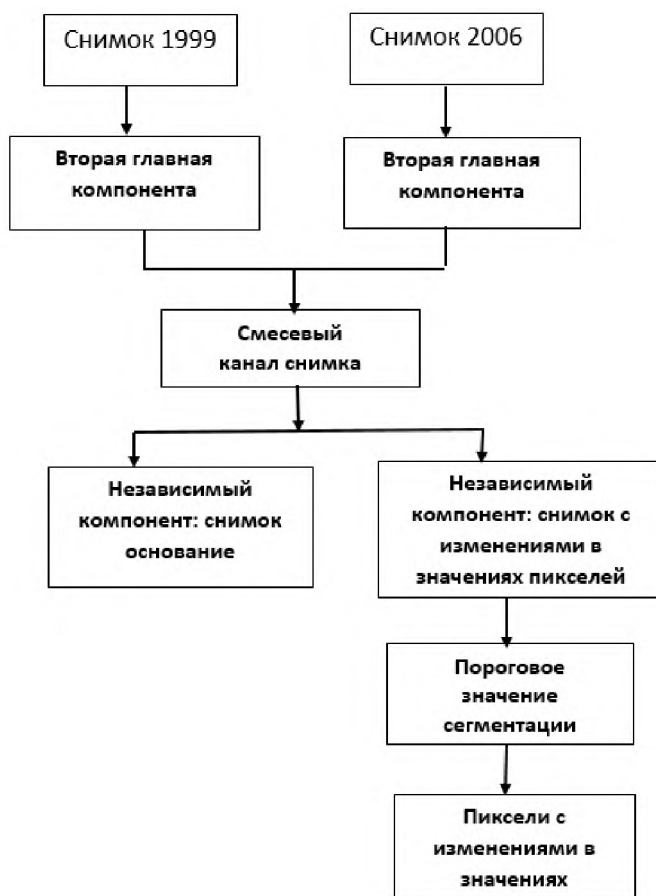


Рис.2 Модель выявления изменений на основе АНК снимка

АНК алгоритм для обнаружения изменений основана на методе пространственной проекции. Для двух случайных величин x_1 и x_2 оценим независимые компоненты для s_1 от x_1 , то можем проецировать x_2 на s_1 . Предположим, что проецируемая величина x_2' , тогда разностное значение величин x_2' и x_2 есть величина изменений x_2 относительно x_1 .

Можно предположить, что подпространство x_1 , как компонента: снимок основание и снимок с изменениями в значениях пикселей в качестве другого компонента. Две компоненты независимы друг от друга.

В результате преобразования исходных снимков за 2001 и 2006 методом главных компонент были получены два снимка. Из них в результате работы АНК алгоритма был получен смесевый снимок (полоса). Который имеет два канала: один из которых снимок

основание, второй снимок с изменениями. Для второго снимка находится пороговое значение. При большом количестве каналов необходимо сжать и проанализировать в каком же из полос (каналов) содержится больше информативной компоненты. Для этого с использованием алгоритма главных компонент были получены компоненты, показанные на рис.3, PCA (Principal Component Analysis) – главные компоненты.

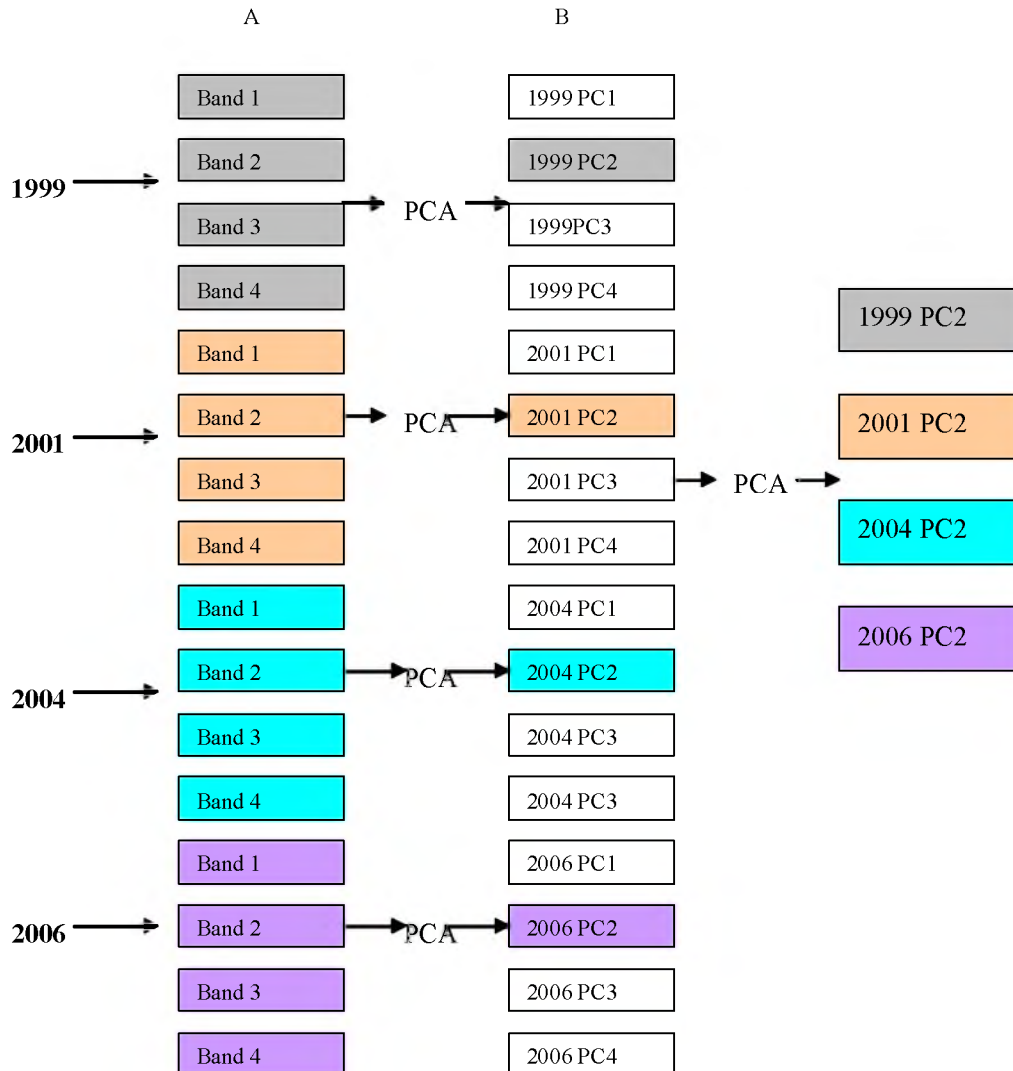


Рис.3 Получение главных компонент (PC)

После проведенных вычислений на основе формулы Байеса было получено процентное содержание компонент, показанных в таблице 1 и обоснован выбор снимков 2001 года и 2006 года для классификации.

Таблица 1.

Даты	PC1(%)	PC2 (%)	PC3 (%)	PC4 (%)
1999PC2	19.66	14.22	14.22	30.69
2001PC2	32.35	36.32	7.68	18.73
2004PC2	10.51	29.36	0.009	50.58
2006PC2	37.48	20.10	58.19	0.008

Областью исследования является окрестные районы Бишкека. Два изображения ЕТМ + получены в 2001 и 2006 годов соответственно и используемыми полосами являются band3,

band4 и band5. Предварительная обработка изображения с выравниванием и очищением от шумов, затем три полосы используются в качестве входных сигналов для преобразования АНК.

Сначала мы использовали алгоритм АНК для классификации изображений 2001 года, чтобы подтвердить осуществимость и эффективность этого алгоритм. Затем использовали указанную выше модель обнаружения изменений для выявления изменений в области между этими двумя годами. Эксперимент результаты показаны следующим образом:



Рис.4 Независимая компонента1
снимок-основание

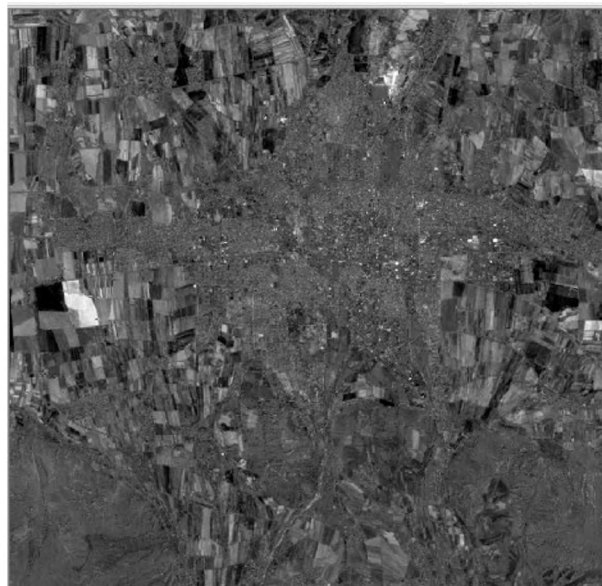


Рис.5 Независимая компонента2
снимок с изменениями

Выводы: Эта статья посвящена анализу независимых компонентов и применение их в классификации и обнаружении изменений в значениях мультиспектральных снимков. Для того чтобы преимущественно извлечь измененные участки на основе АНК, тематическая информация, полученная из независимой компоненты очень полезны для последующей обработки при выявлении и обнаружении изменений. Результаты эксперимента подтвердили целесообразность и эффективность использования АНК в классификации снимков. Тем не менее, в этой статье рассматривается применение АНК, но не сравнивается с другими методами. Метод быстрого АНК используется больше для подготовки данных ДЗЗ и для последующей классификации. Снимки с большим количеством каналов занимают больше времени обработки, причем каналов с полезной информацией может быть сравнительно мало. АНК помогает сжать данные, уменьшить размерность и тем самым вывести на первые места качественную информацию.

Литература:

1. Andrew Ng CS229 Сборник лекций, Глава 12 «Анализа независимых компонент, 2007
2. Васеги С., Жетелова Х. Анализ главных и независимых компонент при обработке снимков, 2015
3. Вершовский Е.А., Разработка методов и алгоритмов кластеризации мультиспектральных данных дистанционного зондирования земли, 2010
4. Кастеллуччо М., «Классификация землепользования в изображениях дистанционного зондирования с помощью сверхточных нейронных сетей», 2015.
5. Кашкин В.Б., Баскова А. А., Рублева Т. В., Власов А. С. «Цифровая обработка аэрокосмических изображений» Версия 1.0 [Электронный ресурс], 2008

6. Колтунов И.А., Великая Я.Г. «Тематическая обработка изображений с помощью смесевых моделей вероятностных моделей», 2016
7. Кочуб Е. В. Анализ методов обработки материалов дистанционного зондирования земли. Белорусский государственный университет, Минск
8. Хювяринен А. «Быстрый и надежный алгоритм анализа независимых компонент», 2011
9. Шовенгердт Р.А. Учебник «Дистанционное зондирование. Методы и модели обработки изображений» Техносфера М.-2010