

Нелинейное снижение размерности гиперспектральных изображений с использованием спектральных углов и учетом пространственного контекста

Е.В. Мясников¹

¹Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34А, Самара, Россия, 443086

Аннотация. В работе предлагается нелинейный метод снижения размерности гиперспектральных изображений. Отличительной особенностью предложенного метода является использование спектральных углов в исходном гиперспектральном пространстве в качестве меры рассогласования между отсчетами изображения, а также учет пространственного контекста пикселей гиперспектрального изображения. Проведенные в работе эксперименты, выполненные на открытом наборе гиперспектральных данных, показали преимущество разработанного метода перед базовыми нелинейными методами снижения размерности и линейным методом главных компонент.

1. Введение

В настоящее время гиперспектральные изображения все более широко используются в самых различных областях деятельности человека. С одной стороны, использование гиперспектральных изображений предоставляет новые возможности, позволяя извлекать информацию о материалах (компонентах), представленных на изображении. С другой стороны, их использование сопряжено с трудностями при решении задач хранения, передачи, обработки и распознавания, обусловленными высокой спектральной размерностью таких изображений. Поэтому важнейшим этапом при использовании гиперспектральных изображений является устранение избыточности таких изображений при сохранении важной спектральной информации.

В связи с этим значительное внимание исследователей уделялось как методам управляемого отбора признаков, так и неуправляемым методам снижения размерности данных. В то время как оба указанных направления неизбежно приводят к некоторой потере спектральной информации, успешные результаты в смысле качества последующей классификации демонстрировались в обоих направлениях.

По ряду объективных причин, таких как большие вычислительные затраты, меньшая устойчивость к изменениям сцены, необходимость наличия классифицированного множества, методы отбора признаков уступают методам снижения размерности. И, хотя при использовании методов снижения размерности данных, исходная интерпретация данных (спектральных сигнатур) оказывается утерянной, последние чаще используются при работе с гиперспектральными изображениями.

Для целей снижения размерности гиперспектральных данных используются как линейные, так и нелинейные методы снижения размерности. Наиболее часто используются линейные методы, включая метод главных компонент (PCA) [1], осуществляющий поиск линейной проекции в подпространство меньшей размерности, максимизирующей разброс данных. Другими линейными методами являются анализ независимых компонент (ICA) [2], метод *projection pursuit* и некоторые другие.

Нелинейные методы снижения размерности, основанные на локально-линейном встраивании (*Locally-linear embedding - LLE*) [3], методе лапласианских собственных карт (*Laplacian eigenmaps - LE*) [4], методе локальных касательных пространств (*Local Tangent Space Alignment - LTSA*) [5], изометрическом встраивании (*ISOMAP*) [6], анализе криволинейных компонент (*Curvilinear component analysis – CCA*), расстояний (*Curvilinear distance analysis – CDA*) [7] и нелинейном отображении (*Nonlinear Mapping - NLM*) [8] используются реже, хотя известно [9], что гиперспектральные изображения дистанционного зондирования подвержены нелинейным эффектам.

Тем не менее, в последние несколько лет популярность нелинейных методов снижения размерности данных непрерывно растет. Было показано, что, в отличие от линейных методов снижения размерности, нелинейные методы позволяют, например, повысить точность классификации подстилающей поверхности и обнаружения объектов. В области визуализации гиперспектральных изображений эти методы также позволяют получить ложноцветовые (псевдоцветовые) представления с желаемыми свойствами.

На сегодняшний день можно указать целый ряд работ, в которых нелинейные методы снижения размерности успешно применялись при обработке мульти- и гиперспектральных изображений. Так, методы CCA и CDA использовались в [10,11] для снижения размерности мультиспектральных изображений. Метод LLE использовался в [12] и позднее использовался вместе с методом LE в работе [13]. Более поздние примеры использования методов LLE, LE и LTSA для анализа гиперспектральных изображений можно найти в работах [14-16], а метода NLM – в работе [17]. В целом, список цитирования, без сомнения, можно расширить.

Во всех указанных нелинейных методах требуется измерять рассогласование между отсчетами гиперспектрального изображения. Наиболее часто используемой мерой для этой цели является евклидово расстояние. Однако существует несколько других мер рассогласования, которые были успешно использованы в анализе гиперспектральных изображений. Наиболее широко используемой среди них является спектральный угол (SAM) [18]. Было показано, что эта мера имеет ряд преимуществ перед евклидовым расстоянием в области анализа гиперспектральных изображений, однако в области нелинейного снижения размерности она использовалась редко.

В частности, можно указать лишь несколько работ [14,19-21], в которых мера SAM использовалась с методами нелинейного снижения размерности. В работе [19] мера SAM использовалась вместе с евклидовым расстоянием на первом этапе метода ISOMAP для определения соседних отсчетов в гиперспектральном пространстве. В работе [20] авторы изучают эффективность меры SAM и евклидова расстояния для снижения размерности с использованием метода лапласианских собственных карт. В работе [14] мера SAM использовалась для улучшения метода лапласианских собственных карт для снижения чувствительности к пропускам данных во временных последовательностях мультиспектральных спутниковых изображений. В работе автора [21] было предложено несколько новых нелинейных методов отображения, основанных на принципе сохранения спектральных углов.

Важным фактом является то, что рассмотренные выше методы снижения размерности действуют в спектральном пространстве. Однако гиперспектральные изображения содержат как спектральную, так и пространственную информацию. И последний тип информации остается неиспользуемым в традиционных методах снижения размерности. Несмотря на то, что мы можем указать множество статей, в которых пространственная информация используется при анализе изображений, в частности, при анализе гиперспектральных изображений (недавний обзор по этой теме может быть найден в [22]), работ, посвященных проблеме использования

пространственной информации гиперспектральных изображений в нелинейных методах снижения размерности, достаточно мало (см., например, [23-25]). В целом большинство используемых для этой цели методов можно отнести к одному из двух способов учета пространственной информации: использование функций ядра, определяемых в пространственно-спектральной области, и использование расширенных мер рассогласования, учитывающих спектральные характеристики локальной окрестности отсчетов изображения.

Таким образом, можно говорить о том, что в целом область применения неевклидовых мер рассогласования совместно с учетом пространственного контекста при нелинейном снижении размерности гиперспектральных данных является на настоящий момент времени недостаточно исследованной.

В настоящей работе предлагается и исследуется нелинейный метод снижения размерности, основанный как на принципе сохранения спектральных углов, так и учитывающий пространственный контекст отсчетов гиперспектрального изображения.

Работа организована следующим образом. В разделе 2 кратко описываются основы базового метода нелинейного отображения; представляется подход, основанный на аппроксимации спектральных углов евклидовыми расстояниями; описывается используемый способ учета пространственного контекста; разрабатывается численная оптимизационная процедура, основанная на методе стохастического градиентного спуска и применимая к гиперспектральным изображениям. В разделе 3 описываются экспериментальные исследования. В этом разделе с использованием гиперспектральных изображений, находящихся в открытом доступе, предложенные методы сравниваются с методом главных компонент (линейным методом) и отдельными нелинейными методами снижения размерности; показывается эффективность предложенного подхода в терминах точности классификации. Работа завершается заключением в разделе 4 и списком используемой литературы.

2. Методы

Перед тем как начать описание предложенного в настоящей работе подхода, приведем основы метода нелинейного отображения.

2.1. Нелинейное отображение

Метод нелинейного отображения относится к классу нелинейных методов снижения размерности, действующих по принципу сохранения попарных расстояний между точками (векторами). Он минимизирует ошибку отображения данных следующего вида:

$$\varepsilon_{ED} = \mu \sum_{i,j=1}^N \rho_{ij} (d(x_i, x_j) - d(y_i, y_j))^2, \quad (1)$$

где N – количество точек данных, x_i – координаты точек в многомерном пространстве R^M , y_i – координаты соответствующих точек в пространстве меньшей размерности R^L , $d(\cdot)$ – функция расстояния, μ и ρ_{ij} – некоторые константы, определяющие конкретную функцию ошибки. Для выражения (1) наиболее часто используются $\mu = 1/\sum_{i<j} d^2(x_i, x_j)$, $\rho_{ij} = 1$, или

$\mu = 1/\sum_{i<j} d(x_i, x_j)$, $\rho_{ij} = (d(x_i, x_j))^{-1}$. В первом случае выражение (1) известно как стресс Крускала [26], а во втором случае то же выражение известно как ошибка отображения Сэммона [8]. Функция расстояния $d(\cdot)$ в выражении (1) обычно является евклидовым расстоянием:

$$d(x_i, x_j) = \|x_i - x_j\| = \sqrt{\sum_{k=1}^K (x_{ik} - x_{jk})^2}. \quad (2)$$

Для минимизации ошибки (1) может использоваться ряд численных методов оптимизации, но наиболее часто используются методы, основанные на алгоритме градиентного спуска [27]. В настоящей работе мы также будем использовать методы этого класса.

Рассмотрим простой алгоритм градиентного спуска. Принимая координаты $Y = (y_1, y_2, \dots, y_N)$ в пространстве низкой размерности в качестве оптимизируемых параметров, в базовой версии

алгоритма производится последовательное уточнение начальной конфигурации точек $Y(0)$ в выходном пространстве с использованием следующего выражения:

$$Y(t+1) = Y(t) - \alpha \nabla \varepsilon, \quad (3)$$

где t – номер итерации, $\nabla \varepsilon$ – градиент целевой функции, в качестве которой здесь выступает ошибка отображения данных (1), α – коэффициент градиентного спуска.

Таким образом, итеративный оптимизационный процесс определяется следующим рекуррентным соотношением для координат точек в выходном пространстве:

$$y_{ik}(t+1) = y_{ik}(t) + 2\alpha\mu \sum_{j=1(j \neq i)}^N \rho_{ij} \frac{d(x_i, x_j) - d(y_i, y_j)}{d(y_i, y_j)} \cdot (y_{ik}(t) - y_{jk}(t)). \quad (4)$$

Это выражение позволяет находить некоторое субоптимальное решение задачи $\varepsilon \rightarrow \min$ путем инициализации выходных координат $y_i(0)$, с последующей итеративной оптимизацией с использованием выражения (4) до тех пор, пока координаты $y_i(t)$ не стабилизируются.

2.2. Нелинейное отображение, основанное на аппроксимации спектральных углов евклидовыми расстояниями

Как указывалось во введении, мера спектрального угла (spectral angle mapper, SAM) [18]

$$\theta(x_i, x_j) = \arccos \left(\frac{x_i \cdot x_j}{\|x_i\| \|x_j\|} \right) \quad (5)$$

часто используется для измерения рассогласования между отсчетами изображения в гиперспектральном пространстве. Наиболее естественным подходом при построении метода нелинейного отображения, действующего по принципу сохранения попарных спектральных углов, будет введение ошибки отображения спектральных углов в форме:

$$\varepsilon_{SAM} = \mu \sum_{i,j=1(i < j)}^N \rho_{ij} (\theta(x_i, x_j) - \theta(y_i, y_j))^2 \quad (6)$$

с последующим применением алгоритма градиентного спуска для получения субоптимального решения. Некоторые варианты реализации такой схемы построения метода нелинейного отображения были апробированы в работе [21]. При этом в указанной работе было также показано, что хорошие результаты при снижении размерности гиперспектральных изображений дистанционного зондирования земли обеспечиваются при аппроксимации исходных мер SAM евклидовыми расстояниями в выходном пространстве. Это наблюдение хорошо согласуется с тем фактом, что расстояние между двумя точками на единичной гиперсфере равно углу между соответствующими векторами. Кроме того, такая схема согласуется с подходом, принятым в многомерном шкалировании. В соответствии с этим подходом мы просто трактуем значения SAM, как некоторые меры рассогласования, принимающие значения из диапазона $[0; \pi]$.

В этом случае выражение (1) принимает вид

$$\varepsilon_{SAM \rightarrow ED} = \mu \sum_{i,j=1(i < j)}^N \rho_{ij} (\theta(x_i, x_j) - d(y_i, y_j))^2 \quad (7)$$

и соответствующий метод оптимизации может быть напрямую получен из выражения (4):

$$y_{ik}(t+1) = y_{ik}(t) + 2\alpha\mu \sum_{j=1(j \neq i)}^N \rho_{ij} \frac{\theta(x_i, x_j) - d(y_i, y_j)}{d(y_i, y_j)} \cdot (y_{ik}(t) - y_{jk}(t)). \quad (8)$$

2.3. Использование пространственного контекста при снижении размерности

Как было сказано во введении, учет пространственного контекста при снижении размерности гиперспектральных изображений может быть выполнен несколькими способами. Так в работе автора [26] были предложены два способа учета пространственного контекста: с использованием функций окна и с использованием порядковых статистик. При этом лучшие результаты в смысле качества последующей классификации были получены при использовании

порядковых статистик. Поэтому в настоящей работе предлагается использовать аналогичный подход применительно к мере SAM. Приведем краткое описание этого подхода.

Рассмотрим выборку a_1, a_2, \dots, a_K . Сортируя выборку и переиндексируя значения выборки так, что $a_{(1)} \leq a_{(2)} \leq \dots \leq a_{(K)}$, получим множество порядковых статистик $a_{(i)}$. 1-ой и K -ой порядковыми статистиками являются $a_{(1)} = \min\{a_i\}$, $a_{(K)} = \max\{a_i\}$, соответственно.

В целях встраивания контекстной информации для каждого отсчета изображения x_i мы рассматриваем пространственное соседство радиуса R , содержащее отсчеты x_1, x_2, \dots, x_K , и упорядочиваем отсчеты окрестности по возрастанию спектрального рассогласования с отсчетом x_i в спектральном пространстве:

$$\theta(x_i, x_{(1)}) \leq \theta(x_i, x_{(2)}) \leq \dots \leq \theta(x_i, x_{(K)}) . \quad (9)$$

Затем признаковое пространство расширяется первыми S порядковыми статистиками (x_i включается в окрестность, как $x_{(1)}$):

$$x_i^* = (x_{(1)}, x_{(2)}, \dots, x_{(S)}) . \quad (9)$$

Далее при снижении размерности в выражениях (7), (8) будем использовать модифицированную меру рассогласования, имеющую вид:

$$\theta^*(x_i^*, x_j^*) = \frac{1}{S} \sum_{s=1}^S w_s \theta(x_{i,s}^*, x_{j,s}^*) . \quad (10)$$

Здесь используется обратное взвешивание $w_s = 1/s$, а S – количество статистик, используемых в качестве пространственного контекста.

2.4. Численная оптимизация, основанная на стохастическом градиентном спуске

К сожалению, описанные выше методы, основанные на простом градиентном спуске, не могут быть напрямую применены к гиперспектральным спутниковым снимкам в связи с высокой вычислительной сложностью и ограничениями по памяти базового метода [27]. Все эти методы требуют выполнения $O(N^2)$ операций на одну итерацию процесса оптимизации, а хранение матрицы предварительно вычисленных попарных рассогласований в многомерном пространстве занимает $N(N-1)/2$ значений с плавающей точкой. По этой причине в настоящей работе будем применять алгоритм стохастического градиентного спуска, основанный на минивыборках.

Для этого алгоритма значение градиента $\nabla \varepsilon$ в выражении (3) оценивается с использованием случайной выборки:

$$\tilde{\nabla} \varepsilon = \sum_{j=1}^R \nabla \varepsilon_{r_j} . \quad (11)$$

где r - случайная выборка (mini-batch), используемая для аппроксимации градиента на итерации t процесса оптимизации, r_j - j -ый элемент этой выборки, R - мощность подмножества r . С использованием этого подхода, мощность минивыборки определяет вычислительную сложность алгоритма на одну итерацию.

Таким образом, для случая гиперспектрального пространства размерности M , редуцированного пространства размерности m , использования S порядковых статистик, количества отсчетов N и подмножеств мощности R вычислительная сложность вычислительная сложность может быть оценена как $O(MmSRN)$.

3. Эксперименты

В экспериментальных исследованиях использовались открытые и хорошо известные гиперспектральные сцены дистанционного зондирования. [28]. В настоящем разделе результаты экспериментов приводятся для сцены Indian Pines (рисунок 1 (а)), 145 x 145 отсчетов, 224 спектральных компоненты, полученных с использованием сенсора AVIRIS. При этом использовалась версия, содержащая только 200 компонент, отобранных путем удаления каналов с высоким уровнем шума и поглощения воды. Эта гиперспектральная сцена

предоставляется вместе с маской истинной классификации (рисунок 1 (б)), которая использовалась для оценки качества классификации.

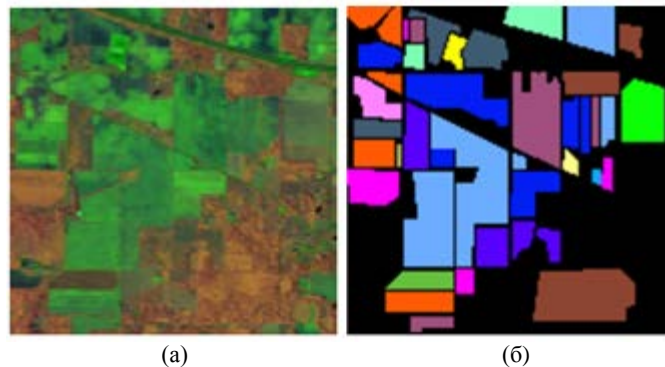


Рисунок 1. Сцена Indian Pines: (а) - представление изображения в псевдо- (ложных) цветах, полученное с использованием нелинейного отображения; (б) - изображение с истинной классификацией (классифицированные отсчеты показаны цветом).

В рамках первого эксперимента оценивалась корректность работы метода снижения размерности. Для оценки качества работы метода использовалась ошибка ε представления гиперспектральных данных в редуцированном пространстве, измеряемая в соответствии с выражением (7), где в качестве меры рассогласования отсчетов в гиперспектральном пространстве использовалась модифицированная мера (10). Пример графика изменения такой ошибки для первых 50 итераций работы метода приведен на рисунке 2 (а). Как видно из приведенного графика, за первые несколько десятков итераций ошибка снизилась на несколько порядков.

На рисунке 2(б) представлен график изменения времени выполнения итерации τ от количества S порядковых статистик. Как видно из рисунка, график линейен, что согласуется с теоретической оценкой сложности, приведенной в разделе 2.4.

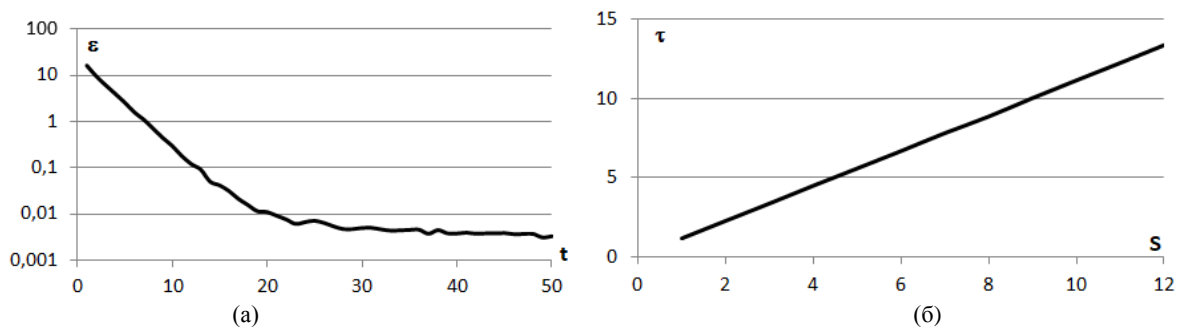


Рисунок 2. Оценка предложенного метода: (а) - зависимость ошибки ε представления гиперспектральных данных в редуцированном пространстве от номера итерации t для размерности выходного пространства $L=5$ и количества порядковых статистик $S=5$; (б) - зависимость времени выполнения итерации τ от количества S порядковых статистик, используемых в качестве пространственного контекста.

Далее в этом разделе оценка предложенного подхода к использованию меры SAM и пространственного контекста выполнена в терминах качества классификации. В частности, использовались два хорошо известных классификатора для признаков, полученных с использованием предложенного метода снижения размерности: классификатор k -ближайших соседей (NN) и машина опорных векторов (SVM). Для выполнения экспериментов все множество классифицированных отсчетов (отсчет с известной истинной классификацией) разделялось на обучающее множество, содержащее 60 процентов выборки, и тестовое подмножество, содержащее 40 процентов выборки. Для оценки качества классификации

использовалась точность классификации, определяемая как доля верно классифицированных отсчетов в общем числе классифицируемых отсчетов.

В рамках второго эксперимента исследовалось влияние количества S используемых в качестве пространственного контекста порядковых статистик на качество классификации. Некоторые результаты проведенного эксперимента показаны на рисунке 3. В целом можно отметить, что почти во всех рассмотренных случаях, чем больше порядковых статистик использовалось, тем выше было качество классификации.

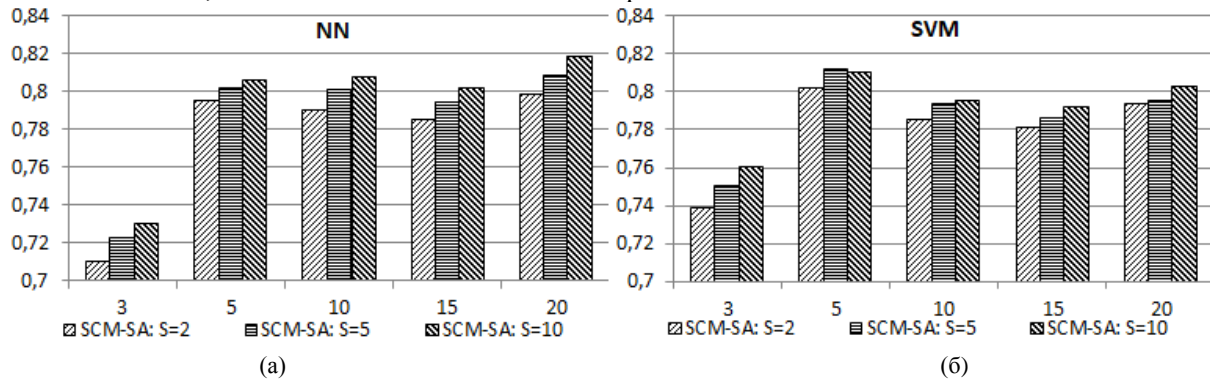


Рисунок 3. Зависимость точности классификации от размерности выходного пространства ($L=3,5,10,15,20$) для различного количества порядковых статистик ($S=2,5,10$).

В рамках третьего эксперимента производилось сравнение предложенного метода с близкими (базовыми) нелинейными методами снижения размерности, а также наиболее широко используемым методом главных компонент. Некоторые результаты проведенного эксперимента показаны на рисунке 4. На указанном рисунке введены следующие обозначения: метод главных компонент - PCA, метод нелинейного отображения, действующий по принципу аппроксимации спектральных углов евклидовыми расстояниями [21] - SAED, метод нелинейного отображения, действующий по принципу сохранения евклидовых расстояний с использованием пространственного контекста [26] - SCM-ED, предложенный метод - SCM-SA.

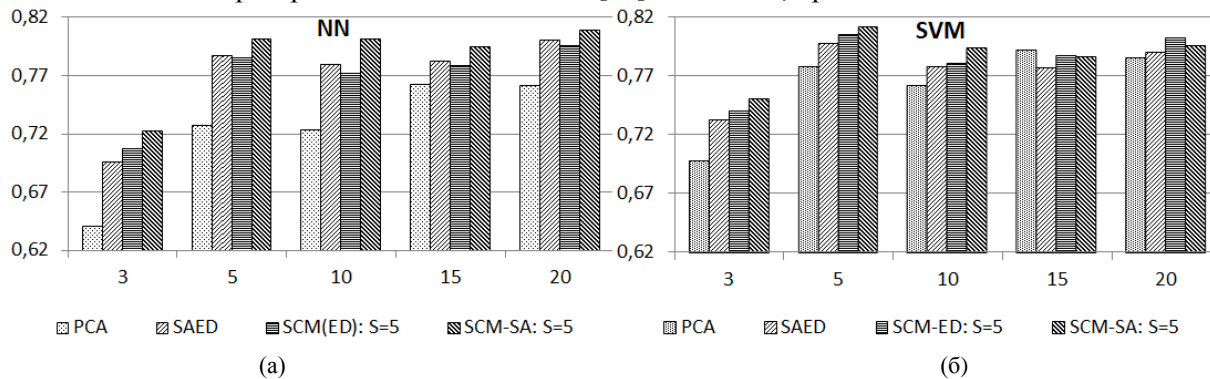


Рисунок 4. Зависимость точности классификации от размерности выходного пространства ($m=3,5,10,15,20$) для различных методов снижения размерности.

Как видно из приведенного рисунка, использование предложенного в работе метода позволило добиться наилучшего качества классификации в большинстве рассмотренных случаев. При этом более заметный эффект был достигнут при использовании классификатора по ближайшему соседу. Это можно объяснить тем фактом, что редуцированное пространство формируется по принципу аппроксимации попарных рассогласований между отчетами в гиперспектральном пространстве евклидовыми расстояниями в редуцированном пространстве.

4. Заключение

В настоящей работе предложен неуправляемый метод нелинейного снижения размерности гиперспектральных изображений, основанный на использовании спектральных углов и учете пространственного контекста отсчетов гиперспектральных изображений. Проведенное в работе экспериментальное исследование показало, что использование предложенного в работе метода позволяет заметно повысить точность классификации отсчетов гиперспектральных изображений.

К недостаткам предложенного метода относится его длительное время работы, линейно возрастающее с количеством порядковых статистик, используемых в качестве пространственного контекста. В связи с этим перспективным направлением исследований представляется сокращение времени работы методов такого класса с использованием современных аппаратных средств параллельных вычислений.

5. Благодарности

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 18-07-01312 А.

6. Литература

- [1] Richards, J. A. Remote Sensing Digital Image Analysis: An Introduction / J.A. Richards, X. Jia, D. E. Ricken, W. Gessner. – Springer-Verlag New York, Inc., 1999.
- [2] Wang, J. Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis / J. Wang, C.-I. Chang // IEEE Trans. Geosci. Remote Sens. – 2006. – Vol. 44 (6). – P. 1586-1600.
- [3] Roweis, S.T. Nonlinear Dimensionality Reduction by Locally Linear Embedding / S.T. Roweis, L.K. Saul // Science. – 2000. – Vol. 290. – P. 2323-2326.
- [4] Belkin, M. Laplacian Eigenmaps and Spectral Techniques for Embedding and Clustering / M. Belkin, P. Niyogi // Advances in Neural Information Processing Systems. – 2001. – Vol. 14. – P. 586-691.
- [5] Zhang, Zh. Principal Manifolds and Nonlinear Dimension Reduction via Local Tangent Space Alignment / Z. Zhang, Z. Hongyuan // SIAM Journal on Scientific Computing. – 2005. – vol. 26(1). – P. 313-338.
- [6] Tenenbaum, J.B. A Global Geometric Framework for Nonlinear Dimensionality Reduction / J.B. Tenenbaum, V. de Silva, J.C. Langford // Science. – 2000. – Vol. 290. – P. 2319-2323.
- [7] Demartines, P. Curvilinear Component Analysis: A Self-Organizing Neural Network for Nonlinear Mapping of Data Sets / P. Demartines, J. Héroult // IEEE Transactions on Neural Networks. – 1997. – Vol. 8(1). – P. 148-154.
- [8] Sammon, J. A nonlinear mapping for data structure analysis / J.W. Sammon Jr. // IEEE Trans. Comput. – 1969. – Vol. 18(5). – P. 401-409.
- [9] Bachmann, C.M. Improved Manifold Coordinate Representations of Large-Scale Hyperspectral Scenes / C.M. Bachmann, T.L. Ainsworth, R.A. Fusina // IEEE Trans. Geosci. Remote Sens. – 2006. – Vol. 44 (10). – P. 2786-2803.
- [10] Journaux, L. Nonlinear reduction of multispectral images by curvilinear component analysis: application and optimization / L. Journaux, I. Foucherot, P. Gouton // International Conference on CSIMTA, 2004.
- [11] Lennon, M. Curvilinear component analysis for nonlinear dimensionality reduction of hyperspectral images / M. Lennon, G. Mercier, M. Mouchot, L. Hubert-Moy // Proc. SPIE. – 2002. – Vol. 4541. – P. 157-168.
- [12] Kim, D.H. Hyperspectral image processing using locally linear embedding / D.H. Kim, L.H. Finkel // First International IEEE EMBS Conference on Neural Engineering. – 2003. – P. 316-319.
- [13] Shen-En, Q. A new nonlinear dimensionality reduction method with application to hyperspectral image analysis / Q. Shen-En, C. Guangyi // IEEE International Geoscience and Remote Sensing Symposium. – 2007. – P. 270-273.

- [14] Yan, L. Improved time series land cover classification by missing observation-adaptive nonlinear dimensionality reduction / L. Yan, D.P. Roy // *Remote Sens. Environ.* – 2015. – Vol. 158. – P. 478-491.
- [15] Sun, W. Nonlinear dimensionality reduction via the ENH-LTSA method for hyperspectral image classification / W. Sun et al. // *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.* – 2014. – Vol. 7(2). – Vol. 6419851. – P. 375-388.
- [16] Hong, D.F. Local manifold learning with robust neighbors selection for hyperspectral dimensionality reduction / D.F. Hong, N. Yokoya, X.X. Zhu // *International Geoscience and Remote Sensing Symposium (IGARSS).* – 2016. – Vol. 7729001. – P. 40-43.
- [17] Myasnikov, E.V. Nonlinear mapping methods with adjustable computational complexity for hyperspectral image analysis / E.V. Myasnikov // *Proc. SPIE 9875.* – 2015. – P. 987508-1-987508-6.
- [18] Kruse, F.A. The Spectral Image Processing System (SIPS) interactive visualization and analysis of imaging spectrometer data / F.A. Kruse, J.W. Boardman, A.B. Lefkoff, K.B. Heidebrecht, A.T. Shapiro, P.J. Barloon, A.F.H. Goetz // *Remote Sens. Environ.* – 1993. – Vol. 44. – P. 145-163.
- [19] Bachmann, C.M. Exploiting manifold geometry in hyperspectral imagery / C.M. Bachmann, T.L. Ainsworth, R.A. Fusina // *IEEE Trans. Geosci. Remote Sens.* – 2005. – Vol. 43(3). – P. 441-454.
- [20] Yan, L. Spectral-angle-based Laplacian Eigenmaps for nonlinear dimensionality reduction of hyperspectral imagery / L. Yan, X. Niu // *Photogramm. Eng. Remote Sens.* – 2014. – Vol. 80(9). – P. 849-861.
- [21] Myasnikov, E. Nonlinear Mapping Based on Spectral Angle Preserving Principle for Hyperspectral Image Analysis / E. Myasnikov, M. Felsberg, A. Heyden, N. Krüger (eds) // *Computer Analysis of Images and Patterns. CAIP 2017. Lecture Notes in Computer Science.* – 2017. – LNCS 10425. – P. 416-427.
- [22] Wang, L. A survey of methods incorporating spatial information in image classification and spectral unmixing / L. Wang, C. Shi, C. Diao, W. Ji, D. Yin // *International Journal of Remote Sensing.* – 2016. – Vol. 37(16). – P. 3870-3910.
- [23] Borhani, M. Kernel Multivariate Spectral-Spatial Analysis of Hyperspectral Data / M. Borhani, H. Ghassemian // *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.* – 2015. – Vol. 8(6). – P. 2418-2426.
- [24] Sun, W. Hyperspectral imagery classification using the combination of improved laplacian eigenmaps and improved k-nearest neighbor classifier / W. Sun, C. Liu, W. Li // *Wuhan Daxue Xuebao (Xinxi Kexue Ban). Geomatics and Information Science of Wuhan University.* – 2015. – Vol. 40(9). – P. 1151-1156.
- [25] Myasnikov, E.V. Exploiting spatial context in nonlinear mapping of hyperspectral image data / E.V. Myasnikov, S. Battiato, G. Gallo, R. Schettini, F. Stanco (eds) // *Image Analysis and Processing - ICIAP 2017. Lecture Notes in Computer Science.* – 2017. – LNCS 10485. – 10 p.
- [26] Kruskal, J.B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis / J.B. Kruskal // *Psychometrika.* – 1964. – Vol. 29. – P. 1-27.
- [27] Myasnikov, E.V. Evaluation of stochastic gradient descent methods for nonlinear mapping of hyperspectral data / E.V. Myasnikov A. Campilho, F. Karray (eds) // *Image Analysis and Recognition. ICIAR 2016. Lecture Notes in Computer Science.* – 2016. – LNCS 9730. – P. 276-283.
- [28] Hyperspectral Remote Sensing Scenes. [Электронный ресурс]. – Режим доступа: [http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral Remote Sensing Scenes](http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes) (01.11.2017).

Nonlinear dimensionality reduction of hyperspectral images based on spectral angles and exploiting the spatial context

E.V. Myasnikov¹

¹Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

Abstract. A nonlinear method for the dimensionality reduction of hyperspectral images is proposed in this paper. A special feature of the proposed method is the use of spectral angles in the initial hyperspectral space as a dissimilarity measure between pixels of an image, as well as taking into account the spatial context of the hyperspectral image pixels. A well-known hyperspectral image dataset was used in the experiments. The experiments showed the advantage of the developed method over the basic nonlinear dimensionality reduction methods and the linear principal component analysis technique.

Keywords: Spectral similarity, dimensionality reduction, nonlinear mapping, spectral angle, SAM, spatial context.