

Гиперспектральные изображения как контейнер для встраивания скрытой информации

Д.А. Шапиро
Самарский национальный исследовательский университет им.
академика С.П. Королева
Самара, Россия
david-shapiro@mail.ru

В.А. Федосеев
Самарский национальный исследовательский университет им.
академика С.П. Королева
Самара, Россия
vicanfed@gmail.com

Аннотация — В работе рассматривается возможность использования гиперспектральных изображений (ГСИ) в качестве контейнера для передачи скрытой информации. Анализируется изменение точности классификации ГСИ после встраивания. Рассмотрены различные методы для классификации ГСИ. Анализируется влияние параметров метода встраивания и объёма встраиваемой информации на качество классификации.

Ключевые слова — стеганографическое встраивание, гиперспектральные изображения, тематическая классификация

1. ВВЕДЕНИЕ

Сегодня на орбите работают искусственные спутники Земли, некоторые из которых ведут съемку в сотнях различных спектральных зон. Одной из задач спутникового мониторинга является получение гиперспектральных изображений (ГСИ) [1]. ГСИ можно представить в виде трёхмерного массива данных. Каждый пиксел такого изображения является вектором мощностей излучения для различных длин волн. В настоящее время ГСИ часто используются для классификации подстилающей поверхности [2], т.е. на основе значений, получаемых в различных каналах данных, принимается решение об отнесении пиксела к определённому типу поверхностного материала.

ГСИ может рассматриваться, с одной стороны, как достаточно ёмкий контейнер для скрытой передачи информации, а с другой стороны, как объект, нуждающийся в защите с помощью цифровых водяных знаков (ЦВЗ) от незаконного распространения или изменения. В первом случае необходимо проанализировать, какой максимально возможный объём данных возможно передать без потери качества. При большом объёме дополнительной информации, встраиваемой в ГСИ, точность классификации может ухудшиться. В данной работе рассматривается влияние амплитуды встраивания и объёма встраиваемой информации на качество классификации.

2. ПРОЦЕСС ВСТРАИВАНИЯ И КЛАССИФИКАЦИИ

В качестве модели для встраивания скрытой информации в нашем исследовании использовался метод QIM. К достоинствам данного метода можно отнести визуальную неразличимость встраиваемого ЦВЗ и стойкость к аддитивному гауссовскому шуму. Встраивание происходит поэлементно, т.е. в каждый пиксел независимо друг от друга по формуле (1):

$$C^w(n_1, n_2) = \left[\frac{C(n_1, n_2)}{2\delta} \right] \cdot 2\delta + \delta W(n_1, n_2), \quad (1)$$

где n_1, n_2 – координаты пиксела ГСИ, C – исходное ГСИ, C^w – ГСИ со встроеным ЦВЗ, W – встраиваемый ЦВЗ, δ – параметр, характеризующий шаг квантования при встраивании ЦВЗ, $[]$ – операция взятия целой части.

Процесс извлечения ЦВЗ происходит по формуле (2).

$$W(n_1, n_2) = \left[\frac{C^w(n_1, n_2)}{\delta} \right] \pmod{2}, \quad (2)$$

При отсутствии искажений процедура извлечения обрабатывает корректно для всех точек. Но в рамках данной работы интерес представляет классификация ГСИ. Для чистоты эксперимента процесс классификации реализовывался отдельно для каждого пиксела, независимо друг от друга. В ходе исследования тестировались 4 известные модели классификаторов: метод опорных векторов, дерево решений, случайный лес, градиентный бустинг. В качестве оценки, характеризующей эффективность проводимой классификации, использовалась метрика Accuracy.

3. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

Для реализации экспериментов были выбраны 4 известных ГСИ: Indian_pines, Kennedy Space Center (KSC), Botswana, Pavia University. Обучение моделей классификаторов проходило на 75% выборки, а её тестирование, соответственно, на 25%. В качестве встраиваемой информации использовалась псевдослучайная бинарная последовательность. В таблице 1 приводятся результаты эксперимента – показатели качества работы различных классификаторов.

Таблица 1. Точность классификации исходных ГСИ

	Метод опорных векторов	Дерево решений	Случайный лес	Бустинг
Indian_pines	0,655	0,865	0,734	0,895
KSC	0,83	0,927	0,879	0,939
Botswana	0,892	0,91	0,85	0,921
PaviaU	0,678	0,935	0,888	0,952

Согласно таблице 1, наилучшие результаты показал градиентный бустинг. Поэтому в дальнейших экспериментах приводятся только результаты работы этого алгоритма. В дальнейшем, демонстрация результатов экспериментов будет проводиться на примере 176-канального снимка KSC заболоченной местности.

В рамках второго эксперимента исследовалось влияние параметров встраивания на качество классификации. Для данного эксперимента ЦВЗ встраивался во все каналы ГСИ.

В базовом варианте шаг квантования δ выбирается одинаковым для всех каналов. Однако во втором эксперименте предпринята попытка использования различных значений δ для каждого канала: шаг квантования выбирался пропорционально одному из показателей, отражающих значимость канала для

решения задачи классификации. В качестве таких показателей использовались следующие величины:

- внутриканальная дисперсия,
- корень из внутриканальной дисперсии,
- дисперсия шума канала,
- параметр значимости (feature_importance), рассчитываемый по обученной на ГСИ модели градиентного бустинга.

Для объективного сравнения различных методик выбора δ , множитель показателя значимости подбирался исходя из равенства отклонений по мере MSE между исходным ГСИ и ГСИ со встраиваемой информацией.

На рисунке 1 видно, что наименьшая потеря качества при классификации возникает при использовании в качестве показателя значимости внутриканальной дисперсии. Худший же результат наблюдается при использовании базового варианта, предполагающего постоянство значения δ для всех каналов.

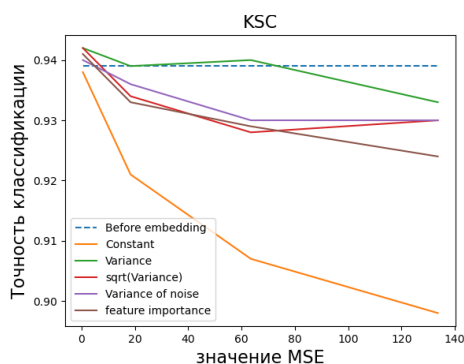


Рис. 1. Точность классификации при различных вариантах выбора параметра δ

В следующем эксперименте проводилось исследование влияния количества каналов, в которые осуществляется встраивание информации, на качество классификации пикселей ГСИ. Для наглядности встраивание осуществлялось при использовании как лучшего, так и худшего варианта выбора δ , выбранных в предыдущем эксперименте. Встраивание происходило во все пиксели с амплитудой встраивания, соответствующей значению $MSE \approx 133,6$.

На рисунке 3 представлен график, показывающий, что при увеличении числа каналов, используемых для встраивания, падает качество классификации. Это связано с тем, что по мере увеличения числа используемых каналов встраиваемая информация затрагивает большее число данных, необходимых для корректной классификации.

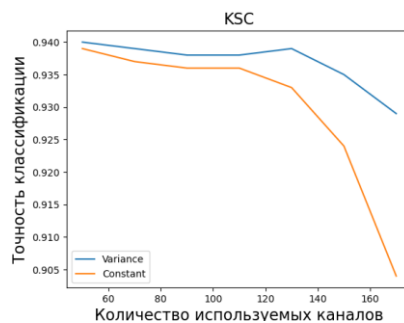


Рис. 2. Зависимость точности от количества заполненных каналов

Наконец, в заключительном эксперименте анализировалась зависимость точности классификации от процента заполнения всех каналов ГСИ встраиваемой информацией. Параметр встраивания δ выбирался равным 13 для всех каналов.

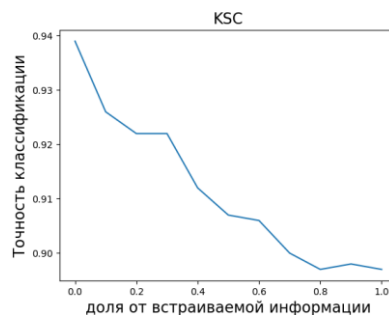


Рис. 3. Зависимость точности классификации ГСИ от доли изменённых пикселей

4. ЗАКЛЮЧЕНИЕ

В данной работе исследуется влияние встраивания скрытой информации в гиперспектральные изображения на качество их классификации при различных параметрах встраивания и различном объёме встраиваемой информации. По итогам экспериментов были сделаны следующие выводы: при реализации встраивания методом QIM наилучшим вариантом выбора параметра δ является задание значения пропорционально дисперсии канала. Наилучшее качество классификации происходит при меньшем объёме передаваемой информации и использовании для передачи лишь порядка 80% наименее значимых каналов.

БЛАГОДАРНОСТИ

Работа выполнена при поддержке РФФИ (проект 19-29-09045).

ЛИТЕРАТУРА

- [1] Шовенгердт, Р.А. Дистанционное зондирование. Модели и методы обработки изображений. / Р.А. Шовенгердт – Москва: Техносфера, 2010. – 560 с.
- [2] Кузнецов, А. В. Сравнение алгоритмов управляемой поэлементной классификации гиперспектральных изображений / А.В. Кузнецов, В.В. Мясников // Компьютерная оптика. – 2014. – Т. 38, № 3. – С. 494–495