

ПРИМЕНЕНИЕ МЕТОДА РСА В ФИЛЬТРАЦИИ ИЗОБРАЖЕНИЙ

Лисименко Л. Д.

*Белорусский государственный университет информатики и радиоэлектроники,
г. Минск, Республика Беларусь*

Научный руководитель: Ролч О. Ч. – канд. техн. наук, доцент, доцент кафедры ПИКС

Аннотация. В данной работе исследован метод машинного обучения без учителя – метод главных компонент (РСА), предназначенный для понижения размерности данных. Экспериментальным путём рассмотрено применение метода РСА к фильтрации изображений. В результате отфильтрован шум в изображениях низкого разрешения

Ключевые слова: метод главных компонент, фильтрация изображений, машинное обучение без учителя

Введение. Подавление шума на изображениях может быть использовано для специализированных целей, например, для увеличения чёткости изображения рентгеновских снимков, снимков, полученных в плохих условиях съёмки, в персонализации и распознавании рукописных знаков.

Основная часть. Метод главных компонент – один из наиболее широко используемых методов машинного обучения без учителя. Его цель – уменьшить количество переменных в наборе данных, сохраняя при этом максимальное количество информации.

Задача машинного обучения без учителя состоит в выяснении зависимости между величинами. В методе главных компонент выполняется количественная оценка этой зависимости путём нахождения списка главных осей координат данных и их использования для описания набора данных. Главные компоненты — это новые переменные, построенные как линейные комбинации или смеси исходных переменных. Эти комбинации выполняются таким образом, что новые переменные не коррелированы, а большая часть информации в исходных переменных сжата или сжата в первые компоненты. При обучении алгоритм определяет главные компоненты и объяснимую дисперсию, характеризующую их информативность. Чем больше дисперсия, тем больше данных может содержать компонента. Использование метода РСА для понижения размерности представляет собой обнуление некоторого числа наименьших главных компонент. Данные по этим осям удаляются и остаются только компоненты с максимальной дисперсией [1, 2].

Исследованные изображения размером 8 на 8 пикселей представляют собой 64-мерные данные. Они дают 64 основные компоненты, но РСА пытается поместить максимально возможную информацию в первую компоненту, затем максимально оставшуюся информацию во вторую и так далее. Идея применения метода РСА для фильтрации зашумленных изображений – восстановление данных с помощью самого крупного подмножества главных (с наибольшей дисперсией) компонент.

Каждая компонента представляет собой набор значений X . Дисперсия рассчитывается по следующей формуле (1):

$$s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{(n-1)} \quad (1)$$

где $(X_i - \bar{X})$ – разность между i -м элементом и средним значением;
 n – количество элементов [3].

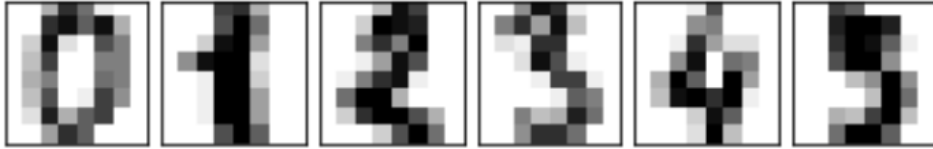


Рисунок 1 – Исходные изображения

На рисунке 1 представлены исходные изображения 8 на 8 пикселей. На рисунке 2 – те же изображения, 42 компоненты каждого из которых содержат добавленный шум. Он должен слабо влиять на компоненты с большой дисперсией. С помощью метода *PCA* произведено обучение с сохранением 50% дисперсии, вычислены главные компоненты и выполнено обратное преобразование. На рисунке 3 представлены отфильтрованные изображения. Положительный результат заметен даже на изображениях низкого разрешения.

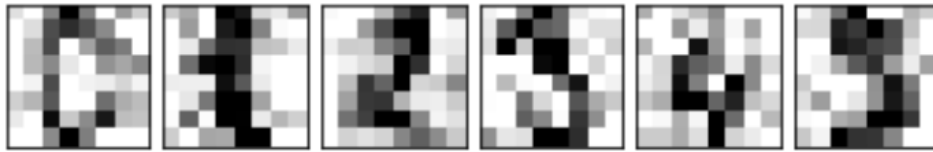


Рисунок 2 – Зашумленные изображения

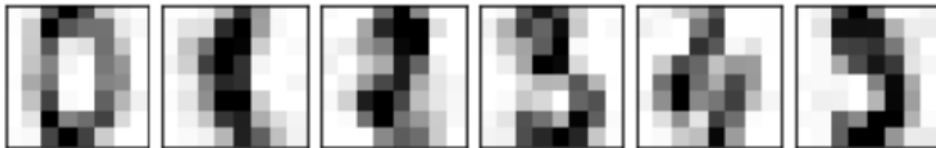


Рисунок 3 – Отфильтрованные изображения

Заключение. В работе применён метод главных компонент для фильтрации изображений. Быстрый метод машинного обучения без учителя может быть использован в восстановлении зашумлённых изображений путём удаления компонент с низкой дисперсией.

Список литературы

1. *A Step-by-Step Explanation of Principal Component Analysis (PCA)* [Электронный ресурс] <https://builtin.com/data-science/step-step-explanation-principal-component-analysis>
2. Плас Вандер Дж. *Python для сложных задач: наука о данных и машинное обучение* / Дж. Вандер Плас – СПб.: Питер, 2018. – 576 с.
3. *A tutorial on Principal Components Analysis* [Электронный ресурс] http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf

UDC 004.048

APPLICATION OF THE PCA METHOD IN IMAGE FILTERING

Lisimenko L. D.

Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

Rolich O.Ch. – PhD, assistant professor, associate professor of the department of ICSD

Annotation. The Principal Component Analysis is an unsupervised machine learning method designed to reduce the dimensionality of data. PCA studied in this work. The applying of this method to image filtering experimentally considered. As a result low-resolution images were denoised

Keywords: principal component analysis, image filtering, unsupervised machine learning