Міністерство освіти і науки України

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «КИЄВО-МОГИЛЯНСЬКА АКАДЕМІЯ»

Факультет інформатики



**Протокол до лабораторної роботи №2**

**З дисципліни „Математичні методи машинного навчання ”**

Виконав

студент 4 курсу

факультету інформатики

Крошин О. А.

1. Завдання лабораторної роботи

1. Сформувати тестову вибірку зображень з вихідного пакету;
2. Провести декомпозицію каналу зеленого кольору тестових зображень з використанням методу головних компонент (PCA):
3. Для каналу зеленого кольору тестових зображень обчислити наступні характеристики розподілу значень яскравості пікселів:
   1. Впорядкувати отримані компоненти вихідного зображення в порядку зменшення значень сингулярних чисел (від найбільшого до найменшого значення);
   2. Провести реконструкцію зображення при використанні лише частки () компонентів розкладу, що характеризуються відмінними від нуля сингулярними числами (). Розглянути випадок, коли змінюється від до з кроком .
   3. Порівняти вихідне та реконструйоване зображення за показником середньоквадратичного відхилення (MSE).
   4. Побудувати графіки залежності , де значення середньо-квадратичного відхилення між вихідним та реконструйованим зображеннями, усереднені по тестовому пакету;
4. Провести моделювання каналу зеленого кольору тестових зображень з використанням Марківських ланцюгів першого порядку :
   1. Отримати стохастичну матрицю для каналу зеленого кольору при обробці пікселів (згідно номеру студента в списку групи, за модулем кількості варіантів):
      1. По горизонталі, зліва направо – ;
      2. По горизонталі, справа наліво – ;
      3. По вертикалі, зверху вниз – ;
      4. По вертикалі, знизу вгору – ;
      5. По головній діагоналі – ;
      6. По головній діагоналі – ;
      7. По побічній діагоналі – ;
      8. По побічній діагоналі – ;
   2. В протоколі роботи графічно показати вид Марківського ланцюга для діапазону яскравості пікселів , де номеру студента в списку групи;
   3. Для отриманих Марківських ланцюгів перевірити виконання властивості регулярності та рекурентності після проходження 5 ітерацій роботи.

2. Порядок виконання роботи та отримані результати

Робота була виконана на мові Python в форматі Jupyter Notebook та за допомогою середовища Jupyter.

1. Для формування вибірки було використано модуль random теки numpy з відповідним зафіксованим сідом (початковим кроком).
2. Для проведення декомпозиції каналу зеленого кольору тестових зображень методом PCA було використано модуль sklearn.decomposition.PCA(), де кількість компонент за замовчуванням відповідає мінімуму(кількість ознак, кількість змінних), а імплементація сингулярного розкладу за замовчуванням – LAPACK full SVD. Масиву компонент відповідає pca.components\_, а сингулярним числам – pca.singular\_values\_

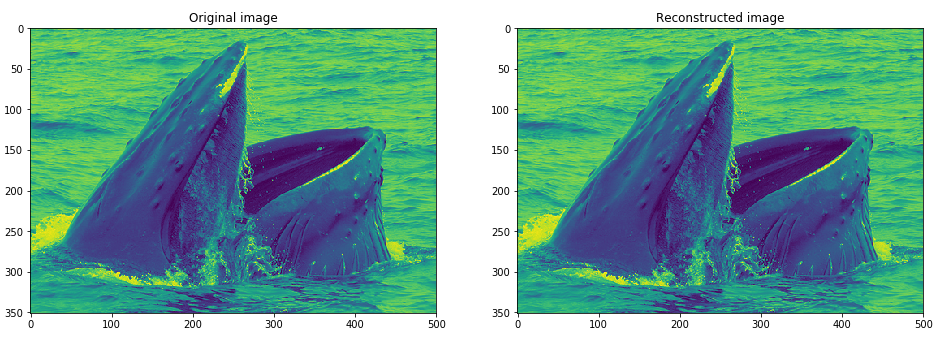
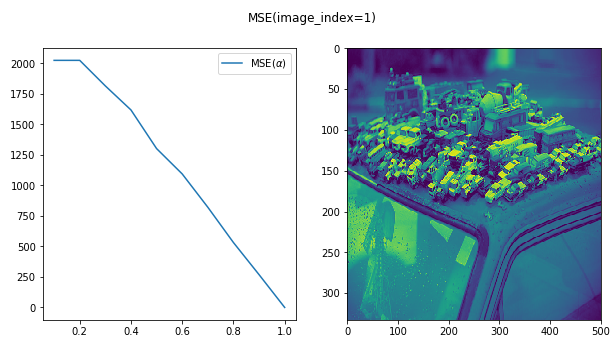


Рис. 1. Порівняння оригінального та відтвореного зображень

Порівнюючи оригінальне та відтворене зображення, бачимо, що візуальна різниця не є суттєвою.

1. Обчислені характеристики розподілу значень яскравості пікселів:

Порівнюючи зображення, відновлене з використанням частки компонентів розкладу (α) бачимо, що середньоквадратичне відхилення між реконструйованим та оригінальним зображеннями зменшується зі збільшенням α. Зауважимо, що для різних зображень графіки суттєво відрізняються, а саме відрізняється рівень α, після якого MSE починає зменшуватись та прямувати до нуля.

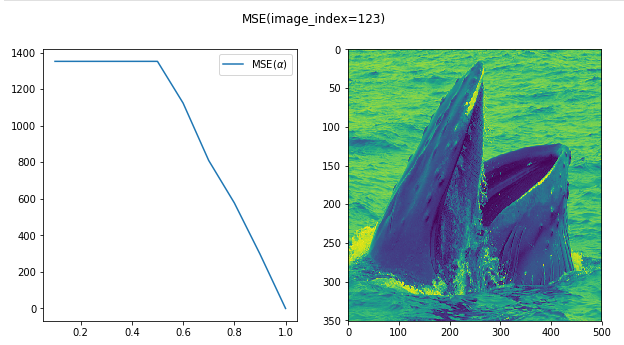


Рисунок 2 Порівняння графіків зміни MSE для різних картинок

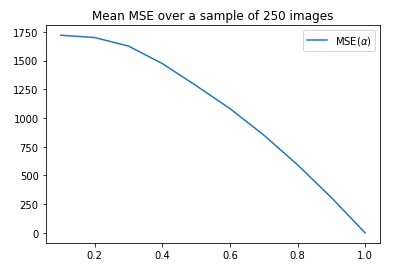
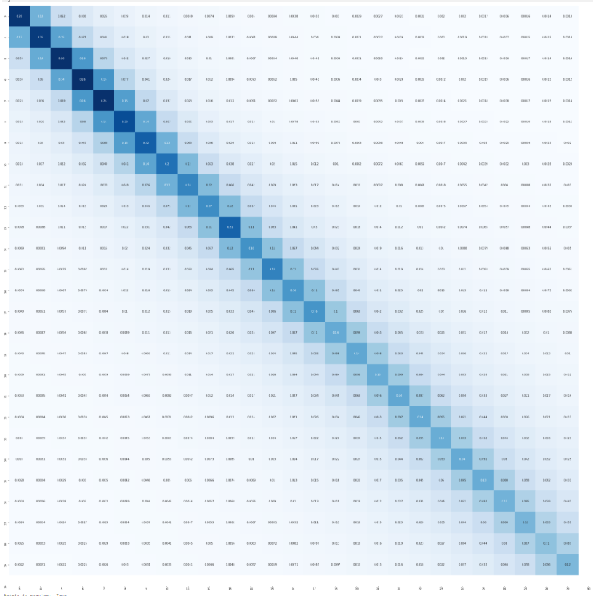
Усереднений по вибірці графік залежності має наступний вигляд:

Рисунок 3 Усереднене за вибіркою MSE

1. Марківські ланцюги першого порядку:

Побудувавши стохастичну матрицю для каналу зеленого кольору при обробці пікселів по вертикалі зверху вниз, ми отримали наступний вид марківського ланцюга для діапазону яскравості пікселів [3;30]:



3. Висновки

В ході лабораторної роботи ми працювали над декомпозицією каналу зеленого кольору тестових зображень з використанням методу головних компонент (PCA). Виконавши декомпозицію зображень зі сформованої вибірки та відтворюючи зображення з використанням різної кількості компонент (в залежності від певного α –кількості компонентів розкладу, використаної для відтворення зображення), ми продемонстрували залежність середньоквадратичного відхилення від α. Використовуючи усереднене середньоквадратичне відхилення по зображеннях вибірки, ми показали, що зі збільшенням α середньоквадратичне відхилення прямує до нуля (рис. 3). Побудувавши марківські ланціюги першого порядку, ми отримали стохастичну матрицю для каналу зеленого кольору при обробці пікселів по вертикалі вниз. Властивості регулярності та рекурентності після проходження 5 ітерацій виконуються