

Отчёт о лабораторной работе №2

Владимиров Эдуард, группа Б05-928

1 ноября 2022 г.

1 Введение

Цель лабораторной работы заключается в применении метода HOSVD для снижения размерности кадров gif-изображения и использовании метода ССМ для обнаружения связи между картинками и звуком. Эксперимент проводится на 15-кадровом черно-белом мультфильме с идущей уткой и сгенерированным звуковым рядом.

2 Постановка задачи

Введём обозначения:

$\underline{X} \in \mathbb{R}^{N \times I \times J}$ - временной ряд кадров: число кадров \cdot размеры изображения

$Y \in \mathbb{R}^N$ - временной ряд звуков

Наша цель заключается в определении наличия связи $\underline{X} \rightarrow Y$. Строгое математическое правило для этого будет предъявлено ниже.

2.1 HOSVD

Мы можем расписать \underline{X} как:

$$\underline{X} \cong \sum_{t=1}^N \sum_{i=1}^{R_i} \sum_{j=1}^{R_j} \sigma_{tij} (u_t^{(1)} \circ u_i^{(2)} \circ u_j^{(3)}),$$

где \circ — это внешнее произведение и

$$U^{(1)} = [u_1^{(1)}, \dots, u_N^{(1)}], U^{(2)} = [u_1^{(2)}, \dots, u_{R_i}^{(2)}], U^{(3)} = [u_1^{(3)}, \dots, u_{R_j}^{(3)}].$$

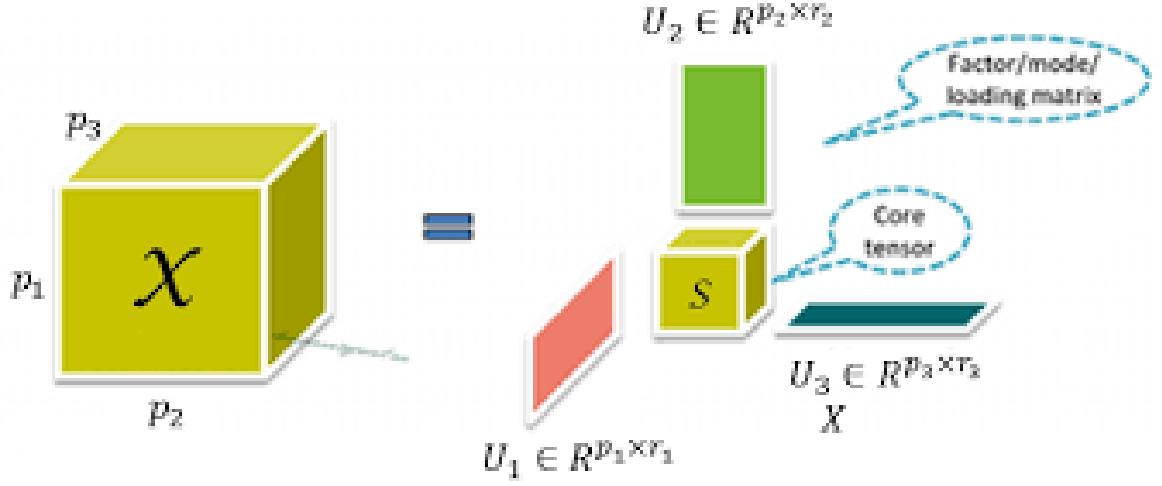


Рис. 1: Иллюстрация метода HOPLS

Или в обозначениях Такера:

$$\begin{aligned} \underline{X} &\cong \underline{G} \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)}, \\ [G]_{tij} &= \sigma_{tij} - \text{core-тензор}, \\ \underline{G} &\in \mathbb{R}^{N \times R_i \times R_j}, \end{aligned}$$

где \times_i — операция матрично-тензорного умножения.

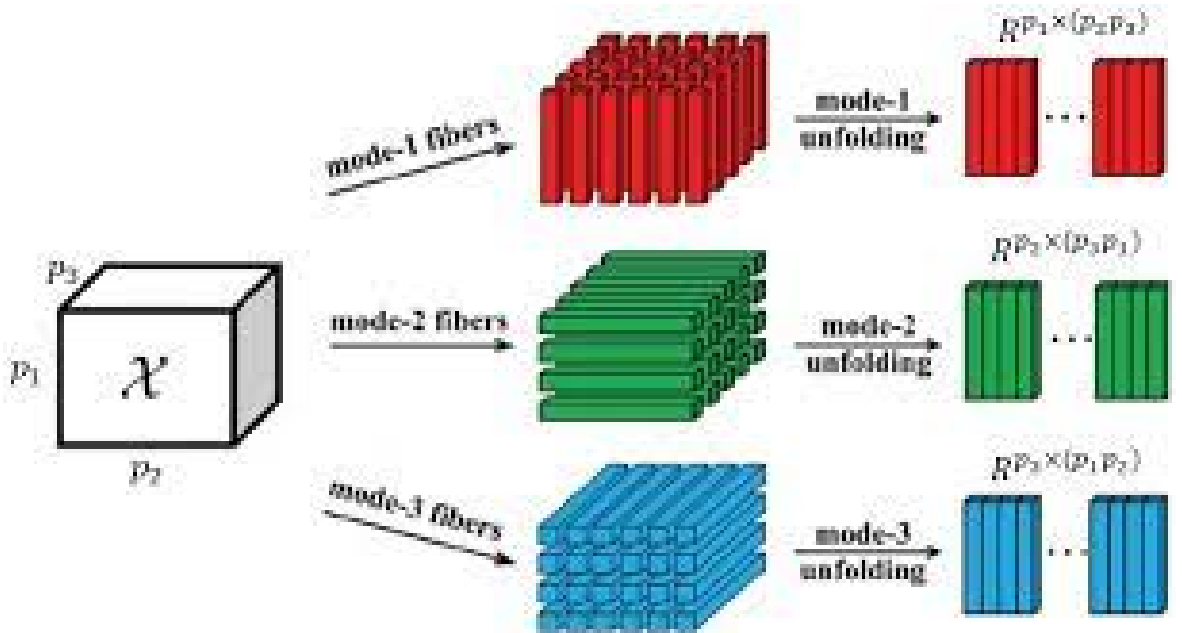


Рис. 2: Матризация тензора \underline{X} , использующая в операция матрично-тензорного умножения \times_l

2.2 CSM

Из аппроксимированного тензора \underline{X} и Y составим их многомерные траекторные матрицы, считая период движения равным T :

$$\mathbf{H}_X \in \mathbb{R}^{(N-T+1) \times T \times I \times J},$$

$$\mathbf{H}_Y \in \mathbb{R}^{(N-T+1) \times T},$$

Определим отображение из траекторного пр-ва звукового сигнала $\mathbb{H}_Y \in \mathbb{R}^T$ в траекторное пространство кадров $\mathbb{H}_X \in \mathbb{R}^{T \times I \times J}$ следующим образом:

$$\varphi : \mathbf{y}_0 \mapsto \hat{\mathbf{x}}_0 = \sum_{i=1}^K w_i \mathbf{x}_i, \quad w_i = \frac{u_i}{\sum_{j=1}^K u_j}, \quad u_i = \exp(-\|\mathbf{y}_0 - \mathbf{y}_i\|).$$

В работе будет рассмотрено два определения наличия связи между временными рядами:

1. Липшицевость отображения φ :

$$\rho_{\mathbb{H}_X}(\varphi(\mathbf{y}_i), \varphi(\mathbf{y}_j)) \leq C \rho_{\mathbb{H}_Y}(\mathbf{y}_i, \mathbf{y}_j) \quad \mathbf{y}_i, \mathbf{y}_j \in \mathbb{H}_Y.$$

2. Высокая корреляция между $\hat{\mathbf{x}}_0$ и \mathbf{x}_0

3 Вычислительный эксперимент

Было взято gif-изображение размером 128×128 и состоящее из 15 кадров. Далее продублировали сигнал 5 раз, чтобы у нас был набор кадров из 5 периодов.

После этого применяется HOSVD с core-тензорами разных размерностей $75 \times \text{hid-size} \times \text{hid-size}$ и восстанавливается аппроксимированный тензор. На рисунке 3 изображены исходное и восстановленные изображения. Из аппроксимированного тензора составляется траекторная матрица картинок. В качестве звукового сигнала Y используется синусоида с периодом в 15 шагов с добавленным нормальным шумом:

$$Y = \left\{ \sin \left(i * \frac{4\pi}{15} \right) + \xi_i \right\}_{i=1}^{75}, \text{ где } \xi_i \sim \mathcal{N}(0, 0.1^2)$$



(a)



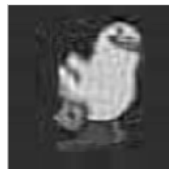
(b)



(c)



(d)



(e)



(f)



(g)



(h)

Рис. 3: Результат применения HOSVD с hid-size, равным (a) 4; (b) 8; (c) 12; (d) 16; (e) 20; (f) 32; (g) 64; (h) 128;

К траекторным матрицам $\mathbf{H}_X, \mathbf{H}_Y$, полученным из аппроксимированного тензора \underline{X} и Y , применяется метод ССМ.

Для каждого $i \in \{47, \dots, 61 = (75 - 15 + 1)\}$ вычисляется $\varphi(\mathbf{H}_Y[i])$. Для проверки на липшицевость в пр-вах \mathbb{H}_X и \mathbb{H}_Y использовались следующие метрики:

$$\rho_{\mathbb{H}_Y}(y_1, y_2) = \|y_1 - y_2\|_2, \quad \rho_{\mathbb{H}_X}(x_1, x_2) = \|x_1 - x_2\|_\infty.$$

Установлено, что при $C = 100$ отображение φ является липшицевым (рисунки 5, 6).

Более того, корреляция между предсказаниями и исходными элементами траекторного пр-ва кадров превосходит 0.74 (рисунок 4).

Таким образом, можно заключить, что искомая причинная связь между временными рядами присутствует.

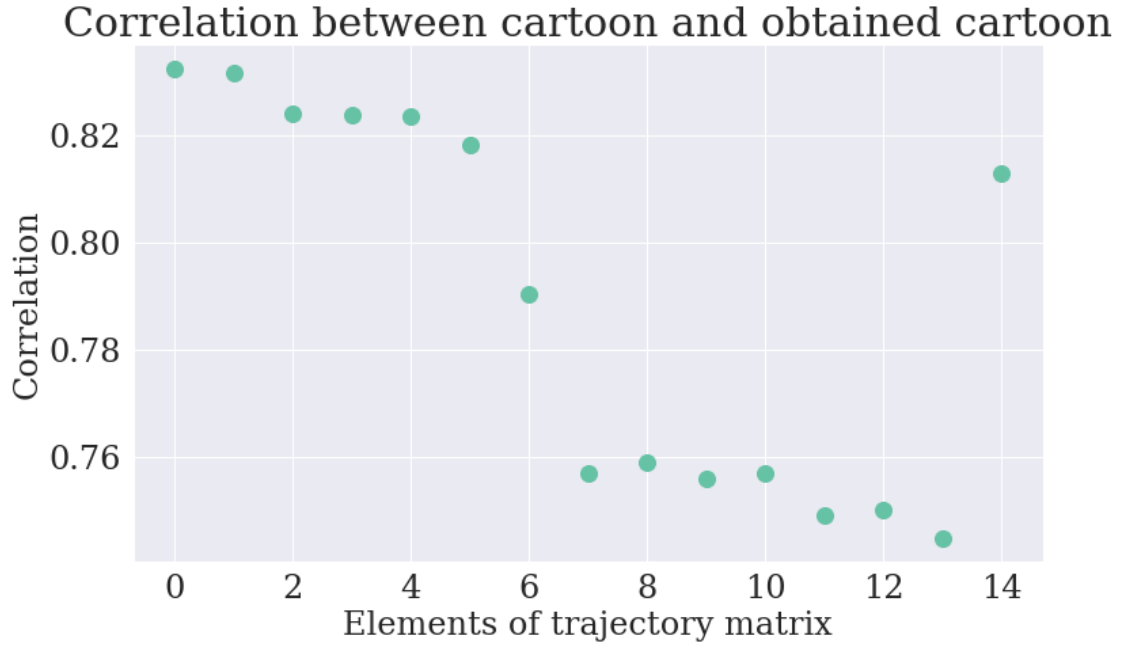


Рис. 4: Корреляция между элементами траекторной матрицы кадров и предсказаниями, полученными с помощью метода ближайших соседей (функции φ)

4 Литература

NULL

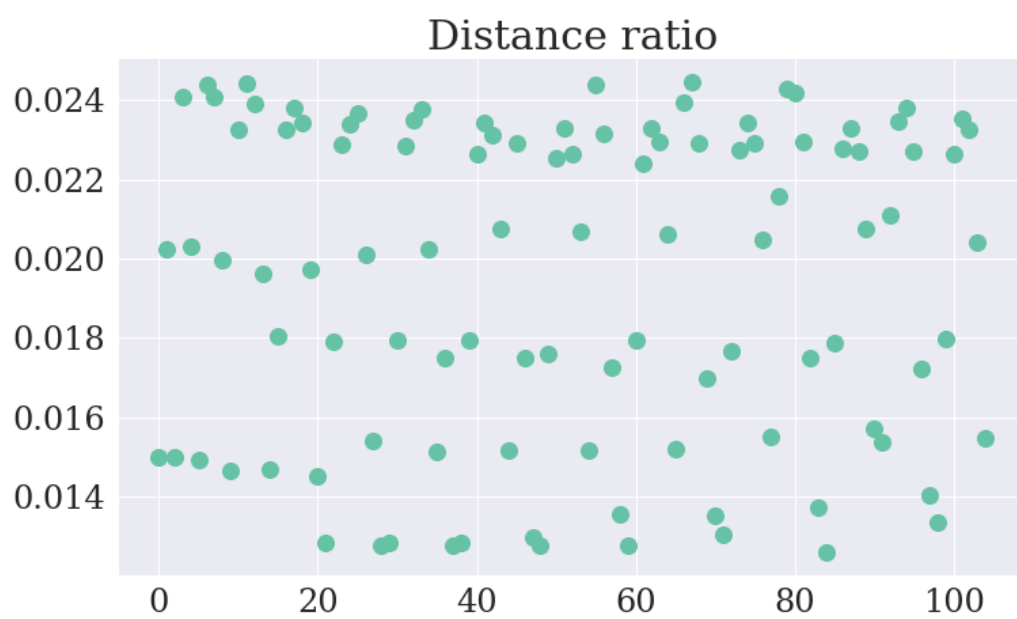


Рис. 5: Отношения расстояний (точечный график)

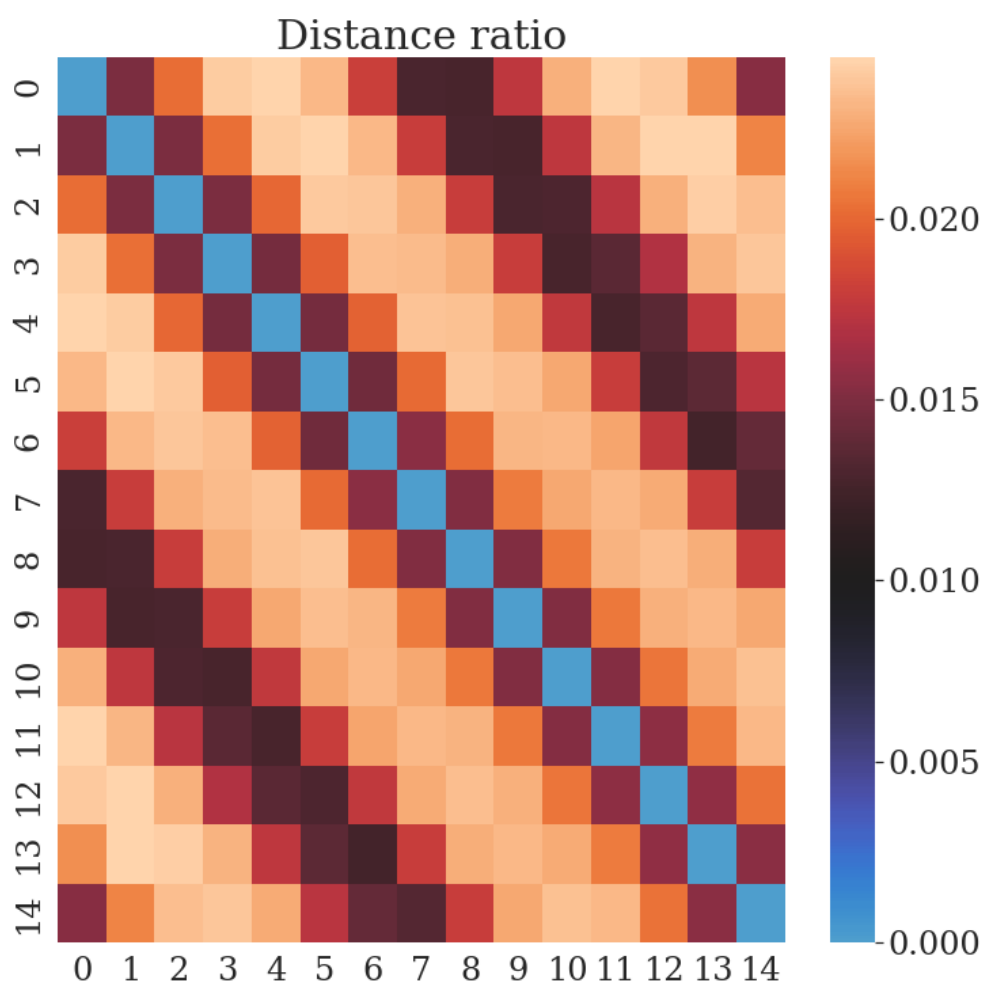


Рис. 6: Отношения расстояний (таблица)