

# Использование сингулярного разложения для оценки информационной емкости компонент изображения

В. Л. Литвинов

Санкт-Петербургский государственный  
электротехнический университет  
«ЛЭТИ» им. В. И. Ульянова (Ленина)  
vlad.litvinov61@gmail.com

Ф. В. Филиппов

Санкт-Петербургский государственный  
университет телекоммуникаций  
им. проф. М. А. Бонч-Бруевича  
9000096@mail.ru

**Аннотация.** Предложена процедура вычисления долей информации, содержащейся в отдельных компонентах сингулярного разложения пиксельной матрицы изображений. Приведена оценка достаточного количества компонент для восстановления изображения с заданной точностью. Результаты работы могут быть использованы для понижения размерности обучающей выборки при распознавании изображений нейронными сетями глубокого обучения.

**Ключевые слова:** количество информации; распознавание образов; сингулярное разложение; SVD-разложение

Будем рассматривать изображения, представленные в виде прямоугольных матриц, элементами которых являются пиксели, заданные вещественными числами в интервале  $[0, 1]$ . Такое представление в точности соответствует формату *jpeg*. Поставим задачу распознавания некоторого изображения  $A$  среди некоторого множества изображений  $\{X\}$ .

Рассмотрим две матрицы  $A = (a_1, a_2, \dots, a_m)$  и  $B = (b_1, b_2, \dots, b_m) \in \{X\}$  как совокупность из  $m$  векторов длины  $n$ . Расстояние между матрицами можно вычислить в евклидовой метрике:

$$\|A - B\| = \sqrt{\sum_{i,j=1}^{n,m} (a_{i,j} - b_{i,j})^2}.$$

В изотропном пространстве такая оценка позволила бы просто решить задачу распознавания. Однако различные компоненты несут далеко не одинаковое количество информации с точки зрения сходства или различия изображений. Поэтому для решения задачи распознавания необходимо найти такой базис в пространстве  $n \times m$ , в котором можно определить линейно независимую систему векторов, которые являются наиболее существенными для распознавания в этом базисе.

Известно, что любую вещественную  $n \times m$  матрицу  $M$  можно представить в виде сингулярного (SVD) разложения:

$$M = UDV^T,$$

где  $U$  – ортогональная матрица размером  $m \times m$ ,  $V$  – ортогональная  $n \times n$  матрица, а матрица  $D$  имеет специальную диагональную форму  $D = \text{diag}(d_1, d_2, \dots, d_r)$ , где  $d_1 \geq d_2 \geq \dots \geq d_r \geq 0$  и  $r$  – ранг матрицы  $M$ .

Таким образом, SVD разложение матрицы  $M$  можно записать в виде:

$$M = UDV^T = \sum_{i=1}^r d_i u_i v_i^T,$$

где  $u_i$  – столбцы матрицы  $U$ , а  $v_i^T$  – столбцы матрицы  $V$ . Столбцы матрицы  $V$  называются правыми сингулярными векторами и всегда ортогональны друг другу, а столбцы матрицы  $U$  называются левыми сингулярными векторами и также ортогональны друг другу.

Сингулярные значения  $d_1, d_2, \dots, d_r$  вычисляются как квадратные корни из ненулевых собственных значений матрицы  $M^T M$ . Эти сингулярные значения можно использовать для поиска подходящей аппроксимации исходной матрицы, поскольку они эквивалентны главным компонентам [1]. Аппроксимируем матрицу  $M$  матрицей  $M^{(q)}$  ранга  $q < r$ :

$$M^{(q)} = \sum_{i=1}^q d_i u_i v_{(i)}^T,$$

у которой линейно независимыми являются лишь  $q$  столбцов. Погрешность аппроксимации принято оценивать на основе форм Фробениуса.

Известно, что сумма квадратов сингулярных значений матрицы  $M$  на самом деле эквивалентна квадрату всего «содержимого» матрицы  $M$  и с ней связана форма Фробениуса, определяемая как

$$\|M\|_F = \sqrt{\sum_{i,j} m_{i,j}^2}.$$

Тогда погрешность аппроксимации составит

$$\sigma = \frac{\|M^{(q)}\|_F}{\|M\|_F},$$

причем, чем больше значение  $q$ , тем лучше  $M^{(q)}$  аппроксимирует  $M$ . Понятно, что после аппроксимации изображение потеряет часть информации.

Рассмотрим насколько полученный результат согласуется с практикой. Возьмем в качестве тестового множества  $\{X\}$  изображения, представленные на рис. 1, соответствующие 25 реальным изображениям в формате *jpeg* размерностью 540 x 720 пикселей. Оценим распределение количества информации на основе сингулярных чисел.

Ниже приведен фрагмент программы на языке R, решающий эту задачу. Для каждого изображения *image[j]* вычисляется сингулярное разложение *r.svd* и определяются доли информации (в процентах), приходящейся на каждую компоненту. Отметим, что общее число сингулярных значений *length(r.svd\$d)* в нашем случае равно 540:

```
for(j in 1:25) {
  # загрузка j-го изображения
  r=readJPEG(image[j])
  # вычисление сингулярного разложения и суммы
  сингулярных значений
  r.svd<-svd(matrix(r, type = "grey"))
  ves = sum(r.svd$d)
  for(i in 1:length(r.svd$d)) {
    # вычисление доли информации в каждой компоненте
    r.svd$d[i] <- (r.svd$d[i]*100)/ves
  }
  for(k in 1:200) {
    # формирование вектора распределения долей
    информации
    distr[k,j]=r.svd$d[k]
  }
}
```

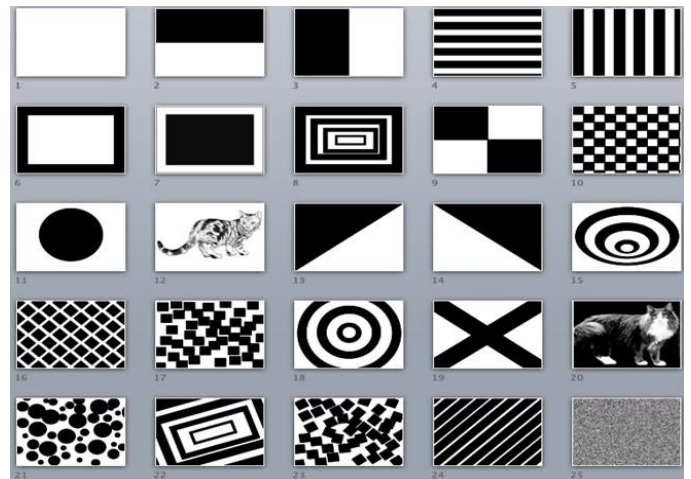


Рис. 1. Тестовое множество изображений

Результирующие распределения долей количества информации по 200 компонентам для каждого изображения приведено на рис. 2. Как видно из рис. 2, тестовое множество изображений упорядочено по значению главной компоненты. Полученное распределение подтверждает факт прямой зависимости плотности распределения от количества значимых компонент [2]. Так, для простейших изображений 1–10 (за исключением 8), достаточно двух компонент для его восстановления с высокой точностью. Для изображения 8 потребуется более 7 компонент.

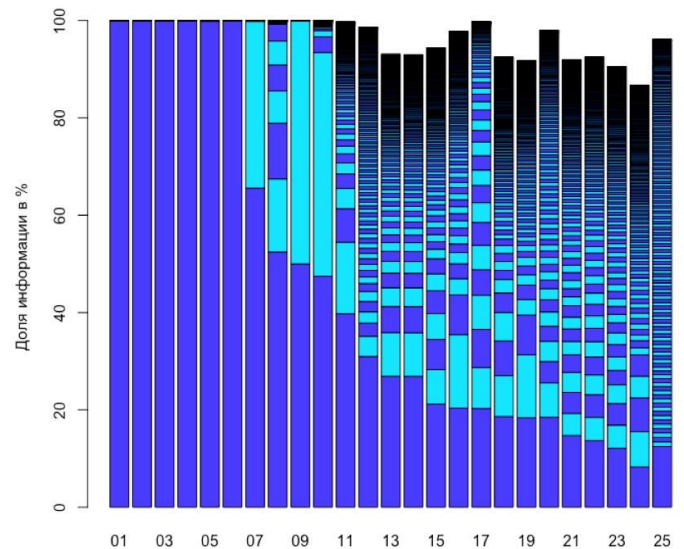


Рис. 2. Распределение информации по компонентам

Для более сложных изображений потребуется большее количество компонент. Так, из рис. 2 следует, что для восстановления ряда изображений потребуется более 200 компонент. Действительно, оценим погрешность аппроксимации  $\sigma$  для числа компонент 2 и 10, используемых при восстановлении изображений 1–25 из тестового множества. Данные расчетов приведены на рис. 3. Если принять допустимой погрешностью

аппроксимации значение равное  $\sigma = 0.98$ , то для десяти изображений будет достаточно всего 2 компоненты, для девяти изображений – 10 компонент.

	k2	k10
1	1.0000000	1.0000000
2	1.0000000	1.0000000
3	1.0000000	1.0000000
4	0.9999999	1.0000000
5	0.9999999	1.0000000
6	1.0000000	1.0000000
7	0.9999991	1.0000000
8	0.9619118	0.9999990
9	0.9999999	1.0000000
10	0.9985978	0.9999981
11	0.9740891	0.9977223
12	0.9734772	0.9880246
13	0.9501275	0.9909310
14	0.9501704	0.9909402
15	0.8854456	0.9852064
16	0.9146289	0.9855852
17	0.8008494	0.9695713
18	0.8677059	0.9855164
19	0.9033177	0.9883094
20	0.8899385	0.9683716
21	0.8281466	0.9536518
22	0.7995026	0.9499380
23	0.7901701	0.9408061
24	0.7114499	0.9522344
25	0.8767750	0.8946060

Рис. 3. Погрешность аппроксимации  $\sigma$  для числа компонент 2 и 10

На рис. 4 приведены случаи распределения долей информации, приходящихся на 10 первых компонент. Показательными здесь являются предельные случаи: белый прямоугольник и «простые фигуры» (изображение 1 – 6) и белый шум (изображение 25), а также некоторый «усредненный» случай (изображение 24).

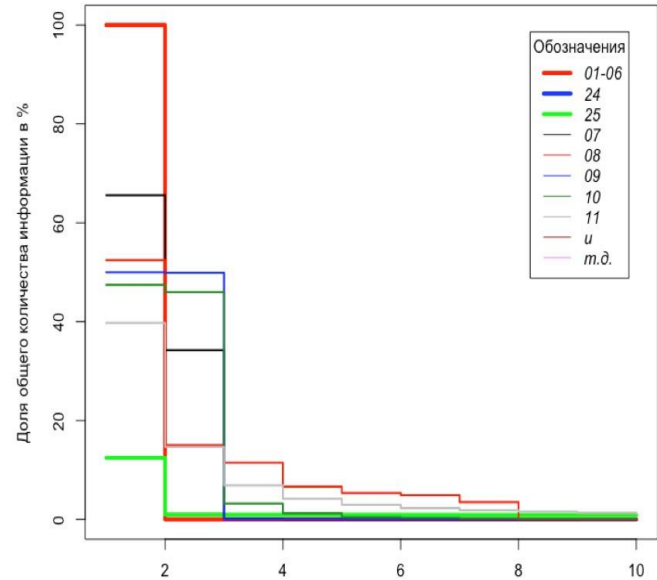


Рис. 4. Доля информации, приходящейся на главные компоненты

Рассмотрение относительного распределения информационных долей (рис. 5), начиная со второй компоненты, наглядно показывает практически линейный спад для белого шума и экспоненциальную зависимость для «усредненного» случая. Понятно, что понятие высокой точности восстановления для белого шума вообще не применимо, поэтому использование только всех компонент позволит восстановить его в деталях.

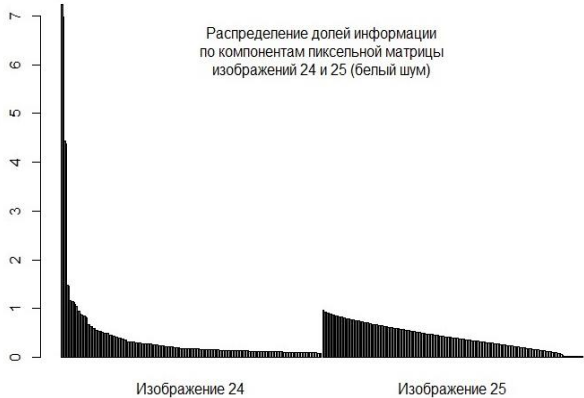


Рис. 5. Относительное распределение информационных долей

### СПИСОК ЛИТЕРАТУРЫ

[1] Губин А.Н., Литвинов В.Л., Филиппов Ф.В. Использование метода главных компонент для распознавания графических объектов // Труды учебных заведений связи. 2016. Т. 2. № 3. С. 27-31.

[2] The use of a matrix decompositions for dimension reduction of training sample / V.L. Litvinov, L.P. Kozlova, F.V. Filippov // 2017 IEEE II International Conference on Control in Technical Systems (CTS) / St. Petersburg, 25-27 oct. 2017. PP. 282-284. DOI: 10.1109/CTS.2017.8109546