# Abstract

Product quantization(PQ) is an effective nearest neighbor search (NNS) method for large-scale high-dimensional data. However, data quantization brings quantization error that may lower the retrieval accuracy. Many methods have been proposed. Among them, the method based on generating optimal PQ codes is very time and memory consuming. To address the problem, we theoretically prove that the more balanced the data volume in each subspace of product quantization is, the smaller the PQ quantization errors. Then an optimal space decomposition based-PQ (OSDPQ) algorithm is proposed. The algorithm solves the optimal space decomposition during product quantization by balancing the data volume in each subspace. Then, we propose the data retrieval method based on the quantization error (DRQE), which can effectively improve the retrieval accuracy of PQ-based NNS methods. Finally, the experimental results show that OSDPQ outperforms NNS methods based on PQ and neural network on 3 datasets. Comparing with the optimized product quantization (OPQ), the memory consumption of our method is reduced by 10%, and the speed of building indexing structure is increased by 10, 4 and 15 times under the close retrieval accuracy. Besides that, we verify the effectiveness of DRQE on PQ-based methods.
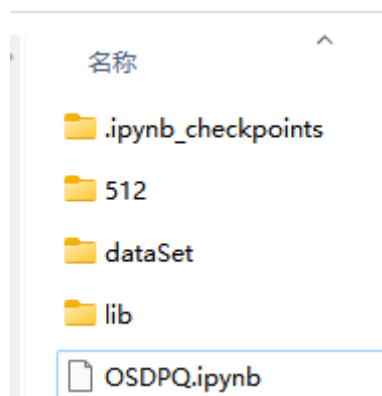
2D graphic
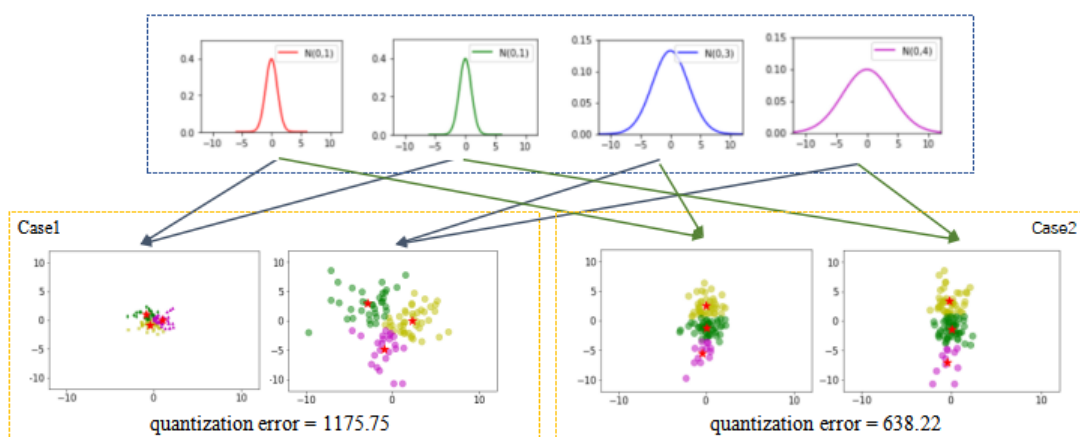
How to use

1. Preparation.

download dataSet: [downlaod link](#) Extraction code: 8888

3. Test

run OSDPQ.ipynb

3. Result

```
############# loading mnist dataset......#############
############# Doing OSDPQ Space Decomposition #################
############# OSDPQ Space Decomposition End ################
################## 开始构建索引 #################
M: 2, Ks: 512, code_dtype: <class 'numpy.uint16'>
M: 2, Ks: 512, code_dtype: <class 'numpy.uint16'>
iter: 20, seed: 123
Training the subspace: 0 / 2
Training the subspace: 1 / 2
Encoding the subspace: 0 / 2
Encoding the subspace: 1 / 2
=========== PQ量化误差为: 234767.19021719877 ===========
====== 第一个子空间的量化误差比例为: 0.4956227074071945 第二个子空间的量化误差比例为: ====== 0.5043772925928055
Encoding the subspace: 0 / 2
Encoding the subspace: 1 / 2
构建索引结构花费了: 37.530853033065796 s
=======================================
```

```
############### Computing Top 1 recall ###############
-------> Search by DRQE <-------
No. First subspace codebook :  3
No. Second subspace codebook :  6
Total Recall   78
Top1 Recall   0.65
=========================================
No. First subspace codebook :  4
No. Second subspace codebook :  6
Total Recall   84
Top1 Recall   0.7
=========================================
No. First subspace codebook :  5
No. Second subspace codebook :  6
Total Recall   87
Top1 Recall   0.725
=========================================
```

Contacts

weilin chen: weierLinC@163.com

shi zhang: shi@fjnu.edu.cn