Codebook Enhancement of VLAD Representation for Visual Recognition

Zhe Wang¹,², Yali Wang¹, Limin Wang¹, Yu Qiao¹
¹ Shenzhen Institute of Advanced Technology, CAS, ²The Chinese University of Hong Kong

Introduction

- **Motivation:**
  - The codebook of VLAD is often limited due to:
    - Poor clustering solution
    - High dimensionality of visual descriptors
    - Global PCA for data pre-processing

- **Contribution:**
  - We propose three approaches for codebook enhancement
    1. Partition of Data
    2. Partition of Feature
    3. Local PCA

Codebook Enhancement and Experiments

We organized the sampled N feature vectors (for codebook generation) into a N × D matrix, where the dimensionality of each feature vector is D. Note that this matrix is the one after PCA pre-processing, where D is the dimensionality that preserves 90% energy of the original feature.

- **Partition of Data**
  - We propose two partition-of-data methods in which k-means is performed with different data partition mechanisms to alleviate poor clustering solution
    1. Hierarchical Clustering (HC)
      - D = D₁ + D₂
      - Clustering the S₁ₙ using D₁ₙ and S₂ₙ using D₂ₙ
      - Every Sₜₜ is constrained by S₂ₙ clusters
    2. Random Clustering (RC)
      - Tr times sample data from codebook matrix
      - Nᵣ clusters are generated by each time in Tr
      - Totally Tr × Nᵣ clusters

- **Partition of Feature**
  - For the N × D matrix, we divide D in three strategies, which means
    1. Non-Overlap (N-OL): D = D₁ + D₂
    2. Partial-Overlap (P-OL): D = D₁ + D₂ + Q₁
    3. Full-Overlap (F-OL): Q₁ = D₁, D = D₂

- **Local PCA**
  - Global PCA: all the clusters share the same dimension-reduction matrix
  - We let each cluster with a dimension-reduction matrix

Comparison to related work

- **Fusion of different methods**
  - For all the experiments, we report Accuracy on HMDB51
  - All of the methods above can be integrated flexibly.

Exploration of setting of different parameters

Accuracy as a function of different parameter settings Nᵣ in our approaches Partition of Dimension (dotted line:Tr=2, full line:Tr=3) and Q₁ in Partition of Feature

Illustration of Our Codebook Generation Methods

- **Comparison to other methods**
  - We compare our best results to related methods on HMDB51 and VOC2007

References