

NMF-KNN: Image Annotation using Weighted Multi-view Non-negative Matrix Factorization

Image annotation/tagging

Superiority

- major drawbacks of other model
 1. require relearning when new images and tags are added to database. Don't support learn online.
 2. The feature fusion is dealt using ad-hoc approaches.
- The ways to address the aforementioned drawbacks
 1. learn query-specific generative model on the features of nearest-neighbors and tags using the proposed NMF-KNN approach which imposes consensus constraint on the coefficient matrices across different features
 2. being query-specific, is unaffected by addition of images and tags in a database

Contributions

- A query-specific model (No global training)
- A natural solution to feature fusion – **coefficient vectors across features to be consistent**
- Handling dataset imbalance through weighted NMF formulation
- $O(n)$ test-time complexity

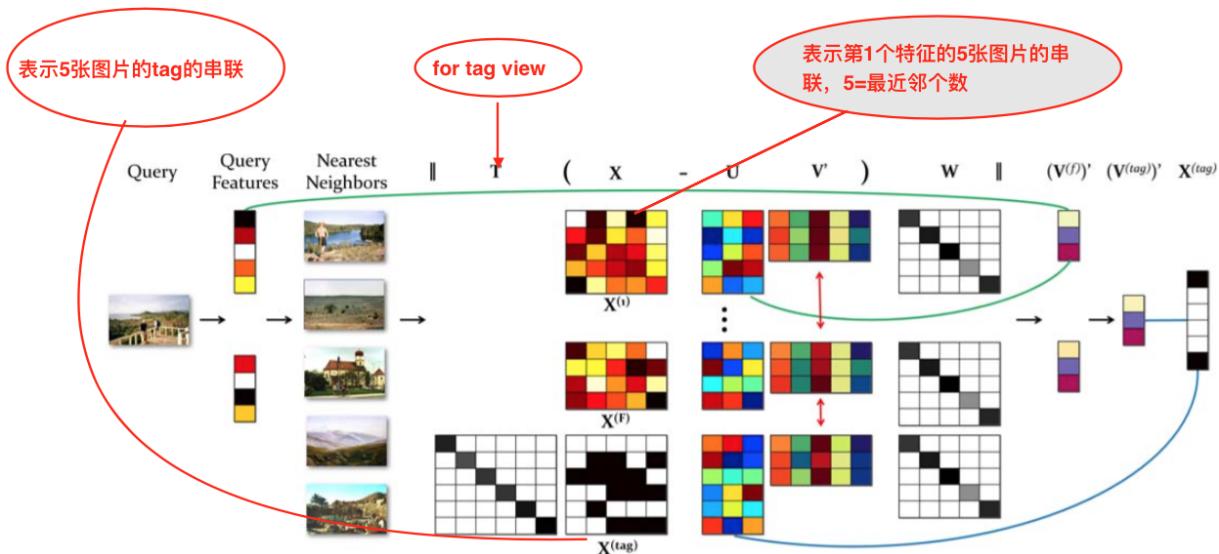
Query-specific Training

$$L = \sum_{f=1}^{F+1} \|T(X^f - U^f V^f)W\|_F^2 + \sum_{f=1}^{F+1} \lambda_f \|W'(V^f Q^f - V^*)\|_F^2$$

- Minimizing the L via an iterative alternative approach
- Training finds the optimum U that minimizes L
- T and W penalizes inaccurate matrix factorization severely for rare tags

Recovering Tags of Query (Testing)

1. **GPSR algorithm** Project query's feature vectors on corresponding basis matrices U – (U get from training)
2. Approximate V^{tag} of query by averaging over F different $V^{visual features}$
3. Predict scores of different tags by computing $U^{tag} \times V^{tag}$



PS. 上图中的 $X^{(tag)}$ 是值得tag的特征，该特征对应的 U 在训练过程中得到，最后通过 $U^{tag} \times V^{tag}$ 得到tag-scores

上图的一个流程就是：

1. 在训练过程中，学习多个特征的映射矩阵U，将各种不同的特征包括tag的特征映射到U上，具体细节是通过NMF实现。通过T和W是用于处理tag的不均衡这个问题上
2. 在test过程中，提取query的多个不同特征，通过这些特征和映射矩阵U得到 V^{tag} ，最后根据训练得到的 U^{tag} 矩阵通过 $U^{tag} \cdot V^{tag}$ 得到tag - scores

Details

- 使用了15种特征， GIST, DenseSIFT, DenseSIFTV3H1, HarrisSIFT, Harris-

SIFTV3H1, DenseHue, DenseHueV3H1, HarrisHue, HarrisHueV3H1

- KNN 用于提取最近邻的图片，提取方式就是讲上述的multiview取平均，作为两两相似度，然后取N张近似图片
- NMF 用于分解feature set – X，变成U和V，因为NMF分解可以很好的保持不同feature之间的一致性，其中U表示的是一个映射矩阵能够实现Query Feature到 $V^{(f)}$ 的映射
- 训练过程中，由于有三个值在变，所以使用固定其中两个值，然后不断迭代的形式来优化求解

annotation的算法分类

1. 基于概率的方法
2. 基于NN的方法
3. 两者结合
4. 近期的热点

流程

该算法的核心就是通过其最近邻的图片来重建测试图片的tag分布，因为我们是在测试图片的近邻图片的基础上进行重建，所以我们要求的是测试图片在近邻图片的基特征U的表示。

首先提取其不同的特征，然后根据这些特征进行聚类，得到最近邻的图片。然后根据不同的特征，将图片进行一个并排，对于每一个特征，做NMF，得到U和V，由于V矩阵具有一定的一致性，所以U矩阵可以很好的表示X的特征变化，然后我们根据测试图片的特征和U矩阵相乘重建出测试图片的V向量。然后根据tag的U矩阵与V相乘得到最后的测试图片的tag分布。T, W用于解决tag的分布不均匀的问题。

下面的公式主要用于NMF，前一部分是X分解为U和V后的误差，后一部分是实现V的一致性条件。求解的过程就是使得L最小化的过程。