

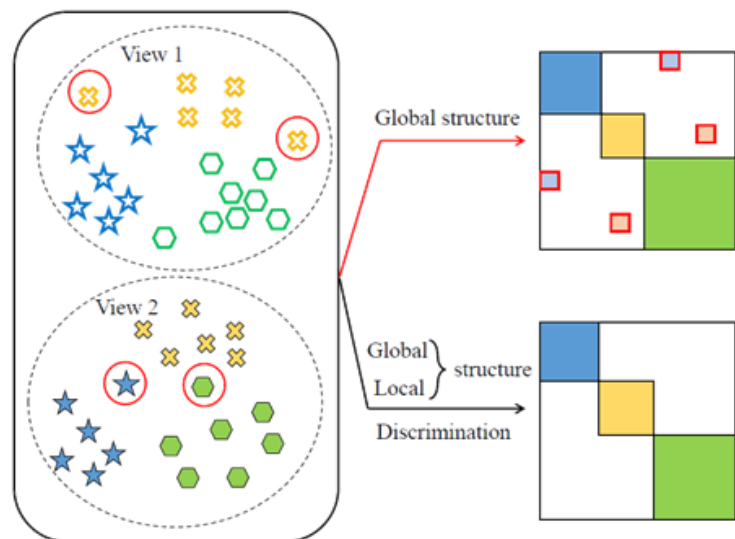
深度多视图子空间聚类

作者：凯鲁嘎吉 - 博客园 <http://www.cnblogs.com/kailugaji/>

1. Deep Multi-view Subspace Clustering with Unified and Discriminative Learning

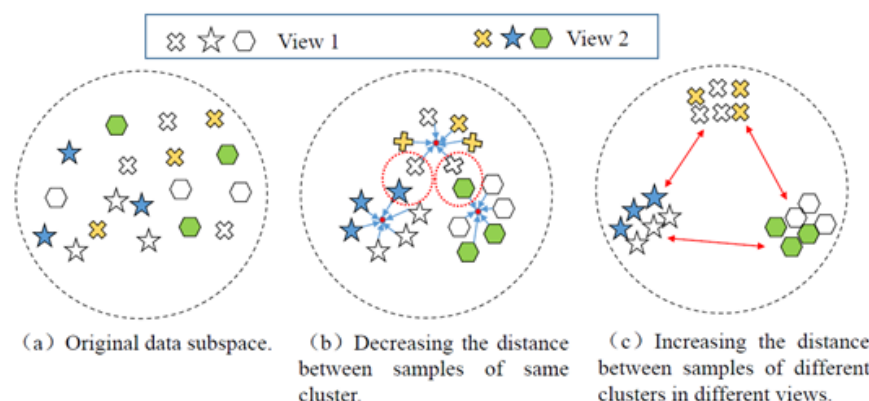
这部分是对Deep Multi-view Subspace Clustering with Unified and Discriminative Learning这篇文章的展开与叙述。与其他多视图聚类方法相比，深度多视图子空间聚类取得了良好的性能。然而，现有的深度多视图子空间聚类只考虑所有视图的全局结构，忽略了每个视图之间的局部几何结构。此外，他们不能学习不同视图的不同簇的判别特征，即簇间差异。为了解决这些问题，本文提出了一种基于统一判别学习的深度多视图子空间聚类算法(DMSC-UDL)。DMSCUDL将全局结构和局部结构与自表示层相结合。全局结构和局部结构相互促进，使同一类样本之间的距离变小。为了在不同视图的不同簇中进一步生成样本，DMSC-UDL在不同视图之间使用了判别约束。这样，DMSC-UDL使得同一簇的样本具有较大的权值，而不同簇的样本具有较小的权值。因此，它可以学习一个更好的多视图聚类共享连接矩阵。大量的实验结果表明，本文提出的多视图聚类方法在性能上优于现有的几种多视图聚类方法。

背景



多视图聚类示意图。左半部分表示多视图样本，右半部分表示共享连接矩阵。视图1和视图2中相同颜色的样本属于同一个簇。连接矩阵的对角线越集中，聚类效果越好。仅对每个视图使用全局自表示结构进行聚类，可能会将权重不准确的红框圈起来的样本聚成错误的聚类。将所有视图的局部结构和判别约束与局部图结构相结合可以获得更好的聚类效果。

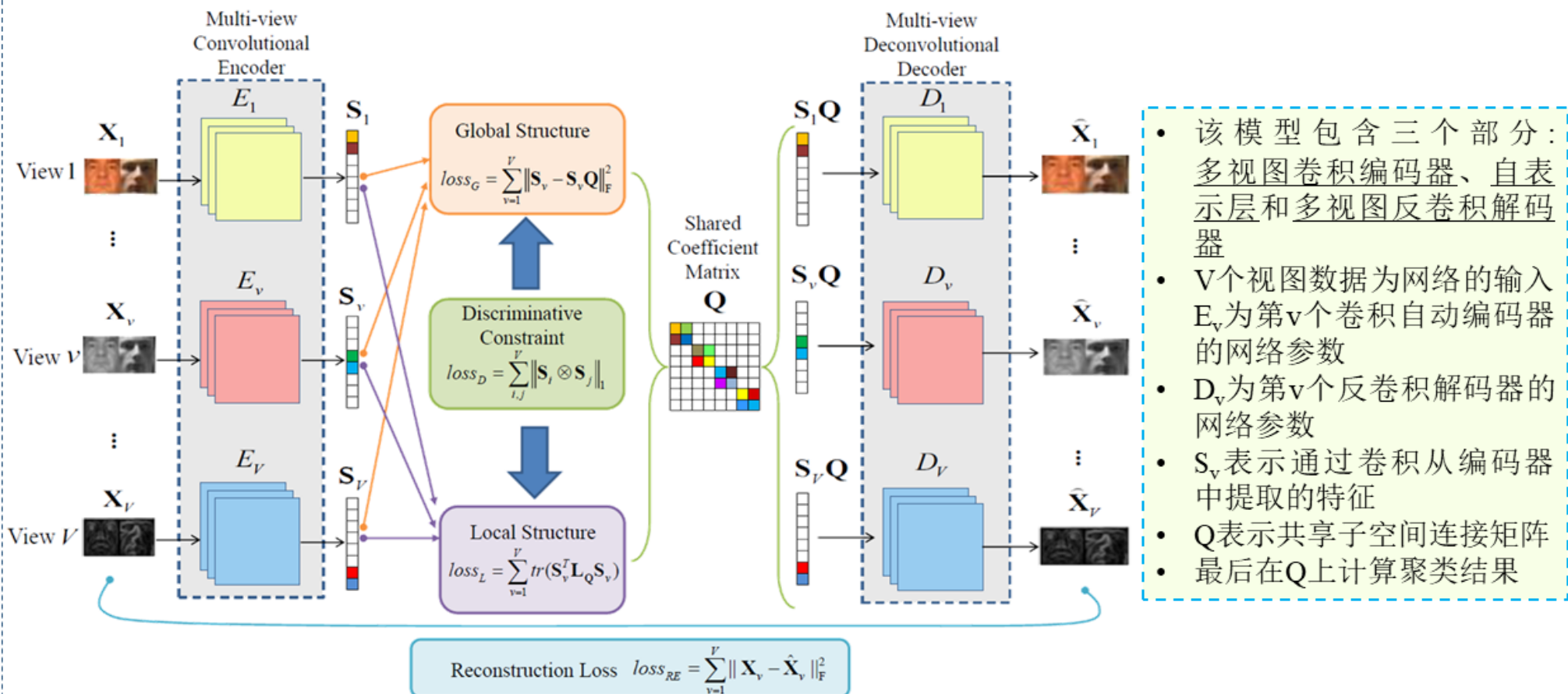
虽然深度多视图子空间聚类方法取得了不错的效果，但它们仍然存在两个局限性：(1)它们只考虑全局结构而忽略局部结构，这对于一致的子空间学习非常重要。因此，只学习全局结构是不全面的。(2)它们只建立了同一聚类样本不同视图之间的紧密关系，而没有考虑不同聚类 and 不同视图之间的区别。也就是说，他们没有考虑属于不同视图的样本在聚类间的差异。



簇间差异示意图。相同形状的实心点和空心点代表不同的视图，不同形状代表不同的簇。

为了克服这些困难，本文提出了一种深度多视图子空间聚类方法——统一判别学习深度多视图子空间聚类(DMSC-UDL)。通过将局部和全局结构学习相结合，DMSC-UDL可以充分利用原始多视图数据的所有信息。此外，在模型中引入了判别约束，以改善不同视图的聚类差异。这样，DMSC-UDL使得相邻样本的权值大，远样本的权值小。因此，DMSC-UDL可以有效地学习一个更好的多视图聚类共享连接矩阵。

► The Structure of DMSC-UDL



Q. Wang, J. Cheng, Q. Gao, G. Zhao and L. Jiao, "Deep Multi-view Subspace Clustering with Unified and Discriminative Learning," IEEE Transactions on Multimedia, doi: 10.1109/TMM.2020.3025666.

DMSC-UDL

$$loss_{All} = loss_{RE} + \lambda_1 loss_G + \lambda_2 loss_L + \lambda_3 loss_D$$



$$\min_Q \underbrace{\sum_{v=1}^V \|X_v - \hat{X}_v\|_F^2}_{\text{Local structure}} + \underbrace{\lambda_1 \left(\sum_{v=1}^V \|S_v - S_v Q\|_F^2 + \|Q\|_F^2 \right)}_{\text{Global structure}} + \underbrace{\lambda_2 \sum_{v=1}^V \text{tr}(S_v^T L_Q S_v)}_{\text{Local structure}} + \underbrace{\lambda_3 \sum_{i,j} \|S_i \otimes S_j\|_1}_{\text{Discriminative constraint}}$$

s.t. $\text{diag}(Q) = 0,$

第一项：重构损失

- 多视图卷积编码器 $X_v \rightarrow S_v$, $S_v = f_v(X_v; E_v)$
- 自表示层 $S_v = S_v Q$, Q 为共享连接矩阵
- 多视图反卷积解码器 $S_v Q \rightarrow \hat{X}_v$, $\hat{X}_v = g_v(S_v Q; D_v)$

第二项：全局结构损失

- 子空间聚类中的自表示层从全局结构上考虑同一聚类样本之间的紧密性。从卷积编码器中提取图像特征 S_v ，从自表达层获得自表达特征 $S_v Q$ 。 S_v 和 $S_v Q$ 之间的损失称为全局结构损失。它使同一类样本之间的距离尽可能接近。

第三项：局部结构损失

- 采用流形学习中的拉普拉斯特征映射(LE)来学习多视图数据的局部结构。
- $L_Q = H_Q - W$, $W = 1/2(|Q|^T + |Q|)$

第四项：判别约束

- 增加对不同视图不同簇的判别
- 使两个向量尽可能正交

约束项：避免 Q 为平凡解(即： Q 的解为单位阵)

► DMSC-UDL

Input:

Multi-view data set $\{\mathbf{X}_v\}_{v=1}^V$, $\lambda_1, \lambda_2, \lambda_3$.

Initial:

learningrate : $lr=0.001$,

optimizer : Adam

epoch = 2000.

1: **While** pre-training not converged **do**:

2: Update **E** and **D** by minimize Eq. (7).

3: Obtain **Q**.

4: **End** pre-training.

5: **While** training not converged **do**:

6: Update **E** and **D** by minimize Eq. (7).

7: Obtain **Q**.

8: **End** training

9: **return** **E** and **D** and **Q**.

Use $\mathbf{W} = \frac{1}{2}(|\mathbf{Q}|^T + |\mathbf{Q}|)$ as the weight of the spectral cluster to do multi-view clustering.

End Procedure

实验过程由三个步骤组成。

第一步: 对网络参数进行预训练

- 利用重构损失对卷积自编码器E和反卷积解码器D进行初始优化。

第二步: 对网络进行正式训练, 得到最终的参数

- 迭代优化卷积自编码器、反卷积译码器和自表示层的参数, 得到最终的参数。
- 首先将通过第一步预训练获得的卷积自编码器参数E和反卷积解码器网络参数D用作网络的初始参数。将多视图数据和常数项参数 λ_1 、 λ_2 和 λ_3 输入到卷积自编码器网络中。
- 经过模型训练, 得到隐层特征和共享连接矩阵Q, 然后将自表示特征输入反卷积解码器, 得到重构的多视图数据。
- 通过最小化总体目标函数来优化E, D和Q。

第三步: 使用Q进行多视图聚类

- 用Q计算W, 然后在W上进行谱聚类, 得到最终的聚类结果。

Q. Wang, J. Cheng, Q. Gao, G. Zhao and L. Jiao, "Deep Multi-view Subspace Clustering with Unified and Discriminative Learning," IEEE Transactions on Multimedia, doi: 10.1109/TMM.2020.3025666.

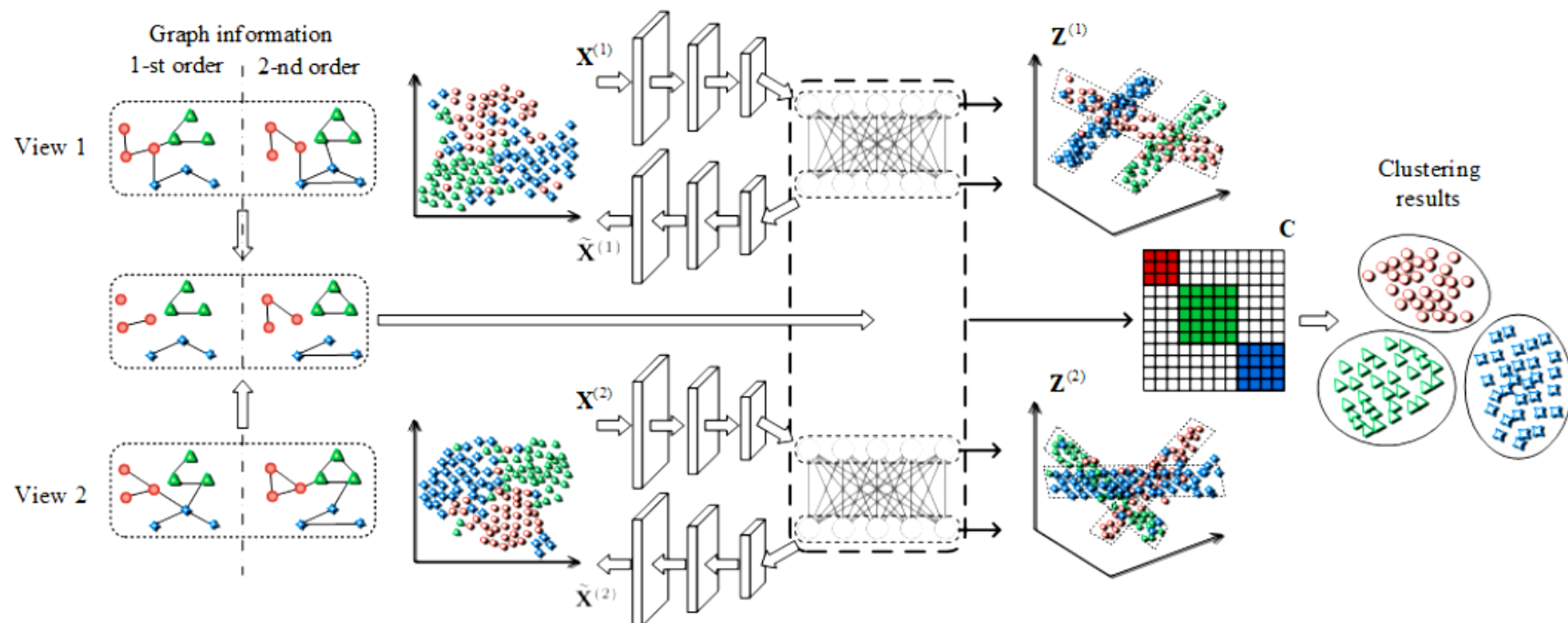
2. Multi-view Subspace Clustering Networks with Local and Global Graph Information

这部分是对Multi-view Subspace Clustering Networks with Local and Global Graph Information这篇文章的展开与叙述。这项研究调查了多视图子空间聚类的问题, 其目的是探索从不同领域或测量收集的数据的底层分组结构。由于数据在许多实际应用中并不总是符合线性子空间模型, 因此大多数现有的基于浅层线性子空间模型的多视图子空间聚类方法在实践中可能会失败。此外, 在大多数多视图子空间聚类方法中, 通常会忽略多视图数据的底层图信息。针对上述局限性, 本文提出了一种新颖的具有局部和全局图信息的多视图子

空间聚类网络，称为MSCNLG。具体而言，在多个视图上使用自编码器网络以实现适合于线性假设的潜在平滑表示。同时，通过将融合的多视图图信息集成到自表示层中，提出的MSCNLG获得了共享的多视图子空间共享表示，可通过采用标准的谱聚类算法获得聚类结果。作为一种端到端的可训练框架，该方法充分研究了多视图的有价值信息。在六个基准数据集上进行的全面实验验证了所提出的MSCNLG的有效性和优越性。

Multi-view Subspace Clustering Networks with Local and Global Graph Information

► The Structure of MSCNLG



本文提出了具有局部和全局图信息的多视图子空间聚类网络(MSCNLG)。如图所示，为了符合线性子空间假设，分别在多个视图上同时进行自编码网络，编码器和解码器之间采用自表示层。自表示层的系数矩阵，即子空间表示，由所有视图共享。该算法利用多视图数据的局部和全局图信息来引导自编码器自表示层中的子空间表示学习过程。如图所示，利用一阶和二阶相似度可以挖掘局部图 and 全局图信息，然后利用共享自表示层约束子空间表示的学习过程。一旦得到所需的子空间表示，就可以使用标准谱聚类算法计算聚类结果。

➤ MSCNLG

$$\mathcal{L} = \frac{1}{2} \sum_{k=1}^v (\underbrace{\left\| \mathbf{X}^{(k)} - \tilde{\mathbf{X}}^{(k)} \right\|_F^2}_{\text{每个视图的重构损失}} + \alpha \underbrace{\left\| \mathbf{Z}^{(k)} - \mathbf{Z}^{(k)} \mathbf{C} \right\|_F^2}_{\text{每个视图的自表示损失}}) + \beta \underbrace{Tr(\mathbf{C}^T \mathbf{L} \mathbf{C})}_{\text{图正则化项, 利用多视图局部和全局图信息来指导多视图子空间表示的学习过程。}}$$

$$= \frac{1}{2} \sum_{k=1}^v \sum_{i=1}^n (\left\| \mathbf{x}_i^{(k)} - \mathbf{f}_i^{(M,k)} \right\|^2 + \alpha \left\| \mathbf{f}_i^{(M/2,k)} - \mathbf{f}_i^{(M/2,k)} \mathbf{C} \right\|^2) + \beta Tr(\mathbf{C}^T \mathbf{L} \mathbf{C})$$

- 第一项：最大限度保留第 k 个视图的信息，以学习潜在的新表示 $\mathbf{Z}^{(k)}$ 。
- 第二项：学习聚类所需的子空间表示 \mathbf{C} 。很明显，应该添加具体约束来指导 \mathbf{C} 的学习过程。 \mathbf{C} ：公共共享系数矩阵
- 第三项：利用多视图数据的底层局部和全局信息来指导 \mathbf{C} 的学习过程。利用不同视图的一阶相似度和二阶相似度来获得融合图。 \mathbf{L} ： \mathbf{W} 的拉普拉斯矩阵。利用Mutual k nearest neighbor (mkNN)计算两个样本点之间的相似性。
- 利用梯度下降法更新迭代得到 \mathbf{C} 之后，在 $1/2(|\mathbf{C}|^T + |\mathbf{C}|)$ 上利用谱聚类得到最终的聚类结果。

➤ Multi-view Graph Fusion

• 第k个视图两个样本之间的一阶相似度

定义为网络中的节点连接关系，用于捕捉局部图结构，并将其简化为完全捕捉图信息进行聚类。

$$\mathbf{W}^{(k)} = \begin{cases} \exp(-\frac{\|\mathbf{x}_i^{(k)} - \mathbf{x}_j^{(k)}\|_2^2}{\sigma^2}), & \text{if } \mathbf{x}_j^{(v)} \text{ and } \mathbf{x}_i^{(v)} \text{ are } mk\text{NN} \\ 0 & , \text{otherwise} \end{cases}$$

• 第k个视图的两个样本之间的二阶相似度

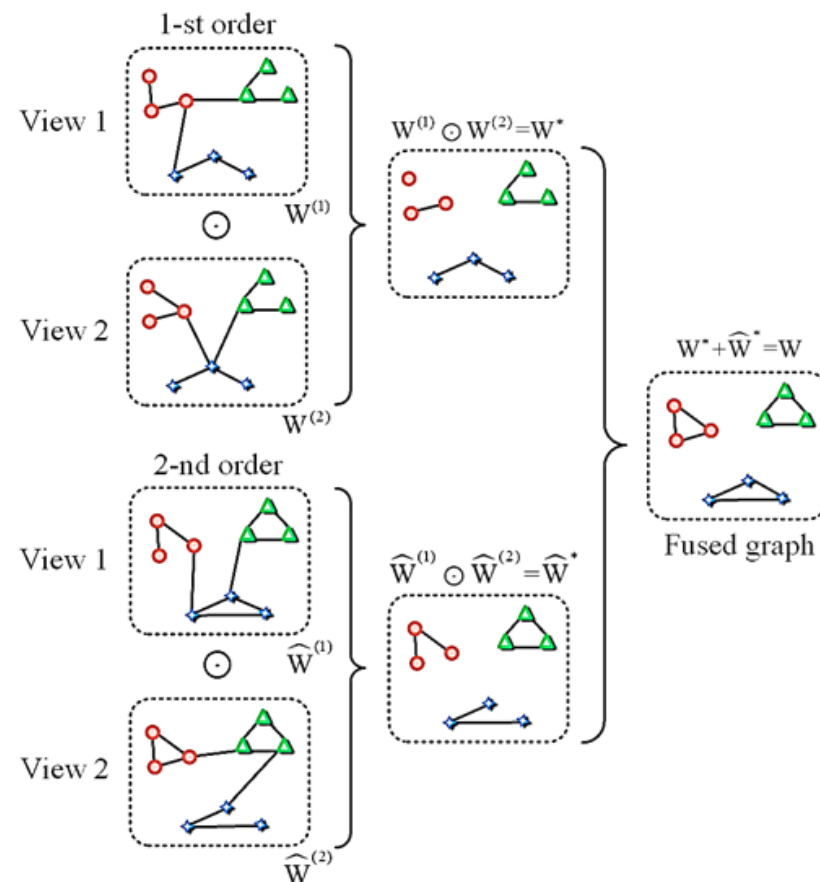
二阶相似度作为一阶相似度的补充，定义为不直接相连的节点的共同邻居节点。根据共享邻域越多的数据点越有可能相似的直觉，二阶相似性由这两个数据点共享的公共邻域的数量决定。显然，多视图的二阶相似度矩阵捕获了全局图信息。

$$\widehat{\mathbf{W}}^{(k)} = \begin{cases} \exp(-\frac{\|\mathbf{w}_i^{(k)} - \mathbf{w}_j^{(k)}\|_2^2}{\sigma^2}), & \text{if } \mathbf{w}_j^{(v)} \text{ and } \mathbf{w}_i^{(v)} \text{ are } mk\text{NN} \\ 0 & , \text{otherwise} \end{cases}$$

• 为整合局部和全局图信息，构造融合多视图图模型

W包含了多视图数据的局部和全局图信息

$$\mathbf{W} = \bigodot_{k=1}^v \mathbf{W}^{(k)} + \bigodot_{k=1}^v \widehat{\mathbf{W}}^{(k)}$$



3. 参考文献

[1] Q. Wang, J. Cheng, Q. Gao, G. Zhao and L. Jiao, "Deep Multi-view Subspace Clustering with Unified and Discriminative Learning," IEEE Transactions on Multimedia, doi: 10.1109/TMM.2020.3025666.

[2] Zhenga Q, J Zhua, Ma Y, et al, "Multi-view Subspace Clustering Networks with Local and Global Graph Information," 2021, Neurocomputing.

Code: <https://github.com/ZQH92/MSCNLG>