Locality Preserving Markovian Transition for Instance Retrieval

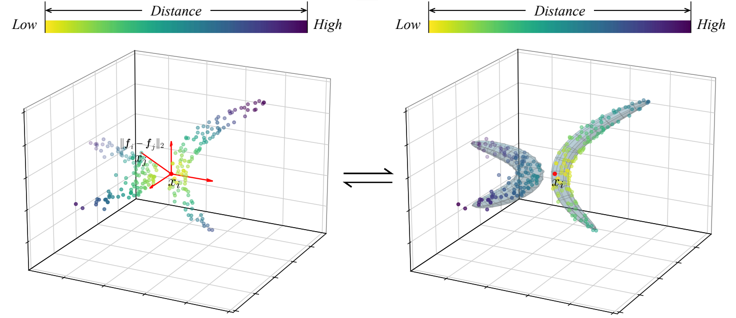
Presentation for the 42nd International Conference on Machine Learning

Jifei Luo[1], Wenzheng Wu[1], Hantao Yao[1], Lu Yu[2], Changsheng Xu[3]

[1]University of Science and Technology of China
[2]Tianjin University of Technology
[3]Institute of Automation, Chinese Academy of Sciences

Table of Contents

Problem Background



(a) Euclidean ranking (b) Manifold ranking

- *Q1: How to effectively model the intrinsic manifold structure?*
- *Q2: How to extract reliable ranking signals from the graph?*
- *Q3: How to mitigate the impact of noise in adjacency graph connections and prevent the loss of essential information?*

Problem Background

In our previous work [LYX24], we proposed a *Bidirectional Similarity Diffusion* strategy to capture the underlying manifold information:

$$\min_{\boldsymbol{F}} \frac{1}{4}\sum_{k=1}^{n}\sum_{i,j=1}^{n} \boldsymbol{W}_{ij}\left(\frac{\boldsymbol{F}_{ki}}{\sqrt{\boldsymbol{D}_{ii}}} - \frac{\boldsymbol{F}_{kj}}{\sqrt{\boldsymbol{D}_{jj}}}\right)^2 + \boldsymbol{W}_{ij}\left(\frac{\boldsymbol{F}_{ik}}{\sqrt{\boldsymbol{D}_{ii}}} - \frac{\boldsymbol{F}_{jk}}{\sqrt{\boldsymbol{D}_{jj}}}\right)^2 + \mu\|\boldsymbol{F} - \boldsymbol{E}\|_F^2$$

where $\boldsymbol{W}$ represents the weighted adjacency matrix, $\boldsymbol{D}$ denotes the diagonal matrix and $\boldsymbol{E}$ is the diffusion source matrix.

Goal: derive the manifold-aware similarity matrix $\boldsymbol{F}$ from $\boldsymbol{W}$.

Bidirectional Collaborative Diffusion

To improve robustness across varying conditions, we develop an adaptive ensemble framework [LWY+25] that integrates multiple similarity matrices:

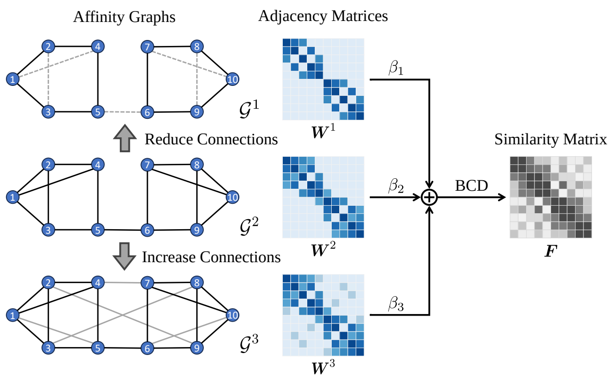


Figure 1: Motivation of Bidirectional Collaborative Diffusion.

Bidirectional Collaborative Diffusion

To improve robustness across varying conditions, we develop an adaptive ensemble framework [LWY+25] that integrates multiple similarity matrices:

$$\begin{aligned} \min_{F} \quad & \sum_{v=1}^{m} \beta_v H^v + \frac{1}{2}\lambda\|\beta\|_2^2 \\ \text{s.t.} \quad & 0 \leq \beta_v \leq 1, \sum_{v=1}^{m} \beta_v = 1 \end{aligned}$$

where H^v is the objective function based on $\boldsymbol{W}^v$ defined as:

$$H^v = \frac{1}{4}\sum_{k=1}^{n}\sum_{i,j=1}^{n} \boldsymbol{W}_{ij}^v\Big(\frac{\boldsymbol{F}_{ki}}{\sqrt{\boldsymbol{D}_{ii}^v}} - \frac{\boldsymbol{F}_{kj}}{\sqrt{\boldsymbol{D}_{jj}^v}}\Big)^2 + \boldsymbol{W}_{ij}^v\Big(\frac{\boldsymbol{F}_{ik}}{\sqrt{\boldsymbol{D}_{ii}^v}} - \frac{\boldsymbol{F}_{jk}}{\sqrt{\boldsymbol{D}_{jj}^v}}\Big)^2 + \mu\|\boldsymbol{F} - \boldsymbol{E}\|_F^2$$

Goal: derive the manifold-aware similarity matrix $\boldsymbol{F}$ from a series of adjacency matrices $\{\boldsymbol{W}^v\}_{v=1}^m$.

Numerical Solution

We propose a numerical solution that decomposes the target function into two sub-problems, enabling systematic iterative approximation.

Optimize $\boldsymbol{F}$ with Fixed β. The optimum is reached via iterative updates:

$$\boldsymbol{F}^{(t+1)} = \frac{1}{2}\sum_{v=1}^{m} \alpha_v \big(\boldsymbol{F}^{(t)}(\boldsymbol{S}^v)^\top + \boldsymbol{S}^v \boldsymbol{F}^{(t)}\big) + (1-\alpha)\boldsymbol{E}$$

where $\alpha_v = \frac{\beta_v}{\mu+1}$, $\alpha = \frac{1}{\mu+1}$ and $\boldsymbol{S} = \boldsymbol{D}^{-1/2}\boldsymbol{W}\boldsymbol{D}^{-1/2}$.

Remark: a higher convergence rate can be achieved with the conjugate gradient method.

Numerical Solution

We propose a numerical solution that decomposes the target function into two sub-problems, enabling systematic iterative approximation.

Optimize $\boldsymbol{F}$ *with Fixed* β. The optimum is reached via iterative updates:

$$\boldsymbol{F}^{(t+1)} = \frac{1}{2}\sum_{v=1}^{m}\alpha_v\big(\boldsymbol{F}^{(t)}(\boldsymbol{S}^v)^\top + \boldsymbol{S}^v\boldsymbol{F}^{(t)}\big) + (1-\alpha)\boldsymbol{E}$$

where $\alpha_v = \frac{\beta_v}{\mu+1}$, $\alpha = \frac{1}{\mu+1}$ and $\boldsymbol{S} = \boldsymbol{D}^{-1/2}\boldsymbol{W}\boldsymbol{D}^{-1/2}$.

Optimize β *with Fixed* $\boldsymbol{F}$. The optimal weight set $\{\beta_v^*\}_{v=1}^m$ can then be obtained by solving the KKT conditions, given by:

$$\beta_v^* = \begin{cases} \dfrac{\sum_{v'\in\mathcal{I}} H^{v'} - |\mathcal{I}|H^v + \lambda}{\lambda|\mathcal{I}|}, v \in \mathcal{I} \\ 0, \quad v \in \{1,2,\ldots,m\}/\mathcal{I} \end{cases}$$

where

$$\mathcal{I} = \big\{v|H^v < (\sum_{v'\in\mathcal{I}} H^{v'} + \lambda)/|\mathcal{I}|, v = 1,2,\ldots,m\big\}$$

Locality State Embedding

For each instance, only the indices belonging to the local region are preserved and assigned the weights by utilizing the corresponding row of $\boldsymbol{F}^*$, resulting in a sparse matrix $\hat{\boldsymbol{P}}$ followed by:

$$\hat{\boldsymbol{P}}_{ij} = \boldsymbol{F}^*_{ij} / \sum_{j \in \mathcal{R}(i,k_1)} \boldsymbol{F}^*_{ij}, \text{ if } j \in \mathcal{R}(i, k_1)$$

Moreover, the local consistency can be further strengthened, yielding the final neighbor-aware probability distribution $\boldsymbol{p}_i$ for each x_i as:

$$\boldsymbol{p}_i = \sum_{j \in \mathcal{N}(i,k_2)} (\kappa \mathbb{1}^{\mathcal{R}}_{ij} + 1) \hat{\boldsymbol{p}}_j / (\kappa |\mathcal{R}(i, k_2)| + k_2),$$

The resulting probability state distributions for each instance in $\mathcal{X}$ can then be organized as $\{\boldsymbol{p}_i\}_{i=1}^{n} \in \mathbb{R}^n$.

Thermodynamic Markovian Transition

To address the problem of unreliable information propagation, we propose a **long-term Thermodynamic Markovian Transition** process to bridge each pair of distributions within the manifold space.

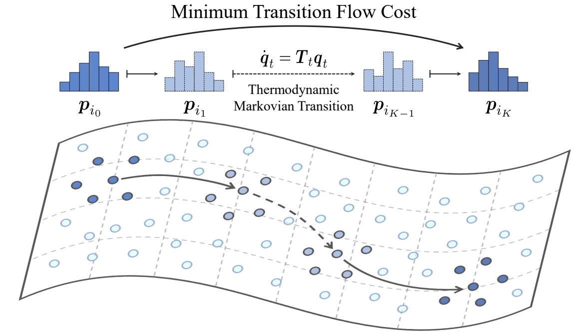


Figure 2: Motivation of the long-term transition process.

Thermodynamic Markovian Transition

To address the problem of unreliable information propagation, we propose a **long-term Thermodynamic Markovian Transition** process to bridge each pair of distributions within the manifold space.
Each step of transition is governed by the **master equation** [VVS23]:

$$\dot{\boldsymbol{q}}_t = \boldsymbol{T}_t \boldsymbol{q}_t$$

We aim to learn a transition policy π that defines a sequence of intermediate states, bridging the source and target distributions within a time interval τ. Formally denoted as $\{\boldsymbol{p}_{i_k}\}_{k=0}^{K} = \pi(\{\boldsymbol{p}_i\}_{i=1}^{n})$, $\boldsymbol{q}_0 = \boldsymbol{p}_{i_0}$, $\boldsymbol{q}_\tau = \boldsymbol{p}_{i_K}$.
We can prove that the minimum transition cost can be computed by:

$$\min_{\pi} \sum_{k=1}^{K-1} \mathcal{W}_1(\boldsymbol{p}_{i_k}, \boldsymbol{p}_{i_{k+1}})$$

Experimental Results

Method	Medium				Hard			
	*R*Oxf	*R*Oxf+1M	*R*Par	*R*Par+1M	*R*Oxf	*R*Oxf+1M	*R*Par	*R*Par+1M
R-GeM (Radenović et al.,2019)	67.3	49.5	80.6	57.4	44.2	25.7	61.5	29.8
AQE (Chum et al., 2007)	72.3	56.7	82.7	61.7	48.9	30.0	65.0	35.9
αQE (Radenović et al.,2019)	69.7	53.1	86.5	65.3	44.8	26.5	71.0	40.2
DQE (Arandjelović & Zisserman, 2012)	70.3	56.7	85.9	66.9	45.9	30.8	69.9	43.2
AQEwD (Gordo et al., 2017)	72.2	56.6	83.2	62.5	48.8	29.8	65.8	36.6
LAttQE (Gordo et al., 2020)	73.4	58.3	86.3	67.3	49.6	31.0	70.6	42.4
ADBA+AQE	72.9	52.4	84.3	59.6	53.5	25.9	68.1	30.4
αDBA+αQE	71.2	55.1	87.5	68.4	50.4	31.7	73.7	45.9
DDBA+DQE	69.2	52.6	85.4	66.6	50.2	29.2	70.1	42.4
ADBAwD+AQEwD	74.1	56.2	84.5	61.5	54.5	31.1	68.6	33.7
LAttDBA+LAttQE	74.0	60.0	87.8	70.5	54.1	36.3	74.1	48.3
DFS (Iscen et al., 2017)	72.9	59.4	89.7	74.0	50.1	34.9	80.4	56.9
RDP (Bai et al., 2019)	75.2	55.0	89.7	70.0	58.8	33.9	77.9	48.0
EIR (Yang et al., 2019)	74.9	61.6	89.7	73.7	52.1	36.9	79.8	56.1
EGT (Chang et al., 2019)	74.7	60.1	87.9	72.6	51.1	36.2	76.6	51.3
CAS (Luo et al., 2024)	80.7	61.6	91.0	75.5	64.8	39.1	80.7	59.7
GSS (Liu et al., 2019)	78.0	61.5	88.9	71.8	60.9	38.4	76.5	50.1
SSR (Shen et al., 2021)	74.2	54.6	82.5	60.0	53.2	29.3	65.6	35.0
CSA (Ouyang et al., 2021)	78.2	61.5	88.2	71.6	59.1	38.2	75.3	51.0
STML (Kim et al., 2022)	74.1	53.5	85.4	68.0	57.1	27.5	70.0	42.9
ConAff (Yu et al., 2023)	74.5	53.9	88.0	61.4	56.4	30.3	73.9	33.6
LPMT	**84.7**	**64.8**	**93.0**	**76.1**	**67.8**	**41.4**	**84.1**	**60.1**

Table 1: Evaluation of the performance on *R*Oxf, *R*Par, *R*Oxf+1M, *R*Par+1M.

Summary & Future Work

- Throughout this line of work, we present a comprehensive framework of *manifold ranking for instance retrieval* in a regime where the extracted features are *noisy and lack sufficient representativeness*.
- Our proposed LPMT framework can be seamlessly integrated into various image retrieval systems as a post-processing module.
- In the future, we aim to extend the application of LPMT to a broader range of *machine learning tasks* and *vector database systems*.

Summary & Future Work

- Throughout this line of work, we present a comprehensive framework of *manifold ranking for instance retrieval* in a regime where the extracted features are *noisy and lack sufficient representativeness*.
- Our proposed LPMT framework can be seamlessly integrated into various image retrieval systems as a post-processing module.
- In the future, we aim to extend the application of LPMT to a broader range of *machine learning tasks* and *vector database systems*.

Thank you for all your attention!

References

[LWY+25] Jifei Luo, Wenzheng Wu, Hantao Yao, Lu Yu, and Changsheng Xu. Locality preserving markovian transition for instance retrieval. In *Proceedings of the 42nd International Conference on Machine Learning*. PMLR, 2025.

[LYX24] Jifei Luo, Hantao Yao, and Changsheng Xu. Cluster-aware similarity diffusion for instance retrieval. In *Proceedings of the 41st International Conference on Machine Learning*. PMLR, 2024.

[VVS23] Tan Van Vu and Keiji Saito. Thermodynamic unification of optimal transport: Thermodynamic uncertainty relation, minimum dissipation, and thermodynamic speed limits. *Phys. Rev. X*, 13:011013, Feb 2023.