QI-IRA: Quantum-Inspired Interactive Ranking Aggregation for Person Re-identification

Chunyu Hu[1]*, Hong Zhang[2,3]*, Chao Liang[1,2,3,4]†, Hao Huang[1]

[1]School of Computer Science, Wuhan University, China
[2]School of Cyber Science and Engineering, Wuhan University, China
[3]National Engineering Research Center for Multimedia Software, Wuhan University, China
[4]Hubei Key Laboratory of Multimedia and Network Communication Engineering, Wuhan University, China
{2019300003061, 2018302180055, cliang, haohuang}@whu.edu.cn

Abstract

Ranking aggregation (RA), the process of aggregating multiple rankings derived from multiple search strategies, has been proved effective in person re-identification (re-ID) because of a single re-ID method can not always achieve consistent superiority for different scenarios. Existing RA research mainly focus on unsupervised and fully-supervised methods. The former lack external supervision to optimize performance, while the latter are costly because of expensive labeling effort required for training. To address the above challenges, this paper proposes a quantum-inspired interactive ranking aggregation (QI-IRA) method, which (1) utilizes quantum theory to interpret and model the generation and aggregation of multiple basic rankings, (2) approximates or even exceeds the performance of fully-supervised RA methods with much less labeling cost, even as low as only two feedbacks per query on Market1501, MARS and DukeMTMC-VideoReID datasets. Comparative experiments conducted on six public re-ID datasets validate the superiority of the proposed QI-IRA method over existing unsupervised, interactive, and fully-supervised RA approaches.

Introduction

Person re-identification (re-ID) aims to retrieve the same person across a group of networking cameras, which has been widely applied in video surveillance and criminal investigation (Wang et al. 2020; Luo et al. 2021; Zahra et al. 2023). The existing re-ID research mainly focuses on designing a single algorithm by improving the feature extraction (Ge et al. 2020; Cho et al. 2022), or metric learning (Yuan et al. 2020; Liu et al. 2022) or re-ranking (Mansouri, Ammar, and Kessentini 2021; Zhang et al. 2022). However, due to the environment complexity, a single re-ID method can not always achieve consistent superiority for different scenarios, which means that it performs better on some datasets and worse on others (as shown in Figure 1). In this case, effectively aggregating the advantages of multiple methods is essential to improve the robustness and generalization of the re-ID methods (Ye et al. 2016).

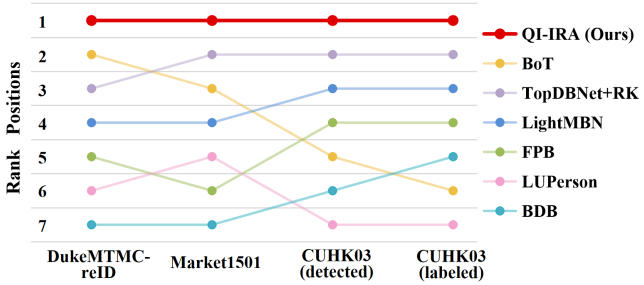


Figure 1: Rank position changes of mAP performance of six representative re-ID methods, including BoT (Luo et al. 2019), Top-DB-Net+RK (Quispe and Pedrini 2021), LightMBN (Herzog et al. 2021), FPB (Zhang et al. 2021), LUPerson (Fu et al. 2021), and BDB (Dai et al. 2019), and their aggregated ranking (by our proposed QI-IRA method) on four public image re-ID datasets.

Ranking aggregation (RA) is a technique for obtaining a better aggregated ranking from multiple basic rankings (Deng et al. 2014), which has been widely applied in social choice (Borda 1781), biometric recognition (Wang et al. 2022), recommendation systems (Bałchanowski and Boryczka 2022) and information retrieval (Wu et al. 2019). The mainstream RA research can be generally classified as unsupervised and fully-supervised methods (Oliveira et al. 2020). Unsupervised methods (Xiao et al. 2021) lack external supervision to directly optimize ranking aggregation performance. Fully-supervised methods (Yu et al. 2020), in contrast, require massive supervisory information which is costly or even unavailable in the practical scenario. Recently, an interactive RA method (IRA) (Huang et al. 2022) has been heuristically proposed to fuse basic rankings with linear weighting guided by relevance feedback. Although IRA improves ranking performance, it doesn't provide the theoretical analysis and explanation of the rationality. And its strategy of updating fusion weights relies on empirically set model parameters, which limits the performance on different datasets (Table 2).

In this paper, we propose a quantum-inspired interactive ranking aggregation method (QI-IRA). The core idea is using quantum theory to model the generation and aggregation

*These authors contributed equally.
†Chao Liang is the corresponding author.

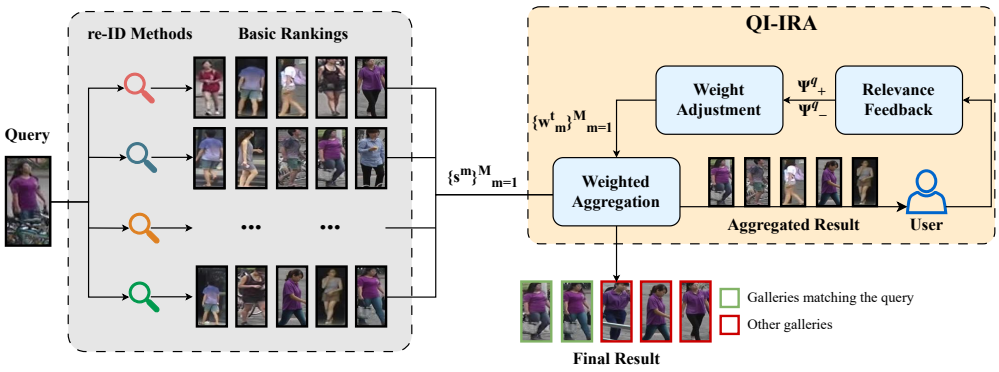


Figure 2: Framework of QI-IRA.

of basic rankings. According to quantum observation theory, each observation corresponds to a set of uniquely orthogonal bases in the state space. When an observation operation is performed on a microscopic particle, its state will collapse into one of the base states with probability (Yanofsky and Mannucci 2008). Hence, the same microscopic particle will produce different results under different observations. If we analogize a microscopic particle to to a gallery image, and various observation operations correspond to different re-ID methods, then *the generation of multiple basic rankings can be explained as different observation results of the same gallery image (microscopic particle) ranked (measured) by various re-ID methods (observation operations).*

Inspired by this, the non-stable performance of a basic re-ID method can be ascribed to its incomplete observation of a complex microscopic particle. Thus, RA is essentially the fusion of multiple incomplete observations into a complete one. In quantum theory, the above fusion process can be formulated as the linear weighting of multiple basic density operators, each corresponding to an in-complete observation. According to Gleason's theorem (Van Rijsbergen 2004), it can be finally transformed into the linear weighting of basic rankings (Eq. 4), providing a theoretical motivation for the linear weighting idea in previous RA methods such as ER (Mohammadi and Rezaei 2020), CSRA (Yu et al. 2020) and IRA. On this basis, QI-IRA introduces a new interactive fusion weight updating scheme by aggregating quantum observation probabilities of feedback samples.

The framework of QI-IRA is shown in Figure 2. First, we obtain basic rankings from multiple re-ID methods, then generate an initial ranking by mean RA and present it to the user. The user conducts relevance feedback to the top-K items in the aggregated result. After that, QI-IRA utilizes the feedback information to adjust the fusion weights of basic rankings, generating the new round RA result. The above process is repeated until the user is satisfied with the current result. Compared with existing RA methods, QI-IRA can achieve higher ranking performance with much less labeling cost, thus effectively address the native dilemma between performance and cost in practical applications.

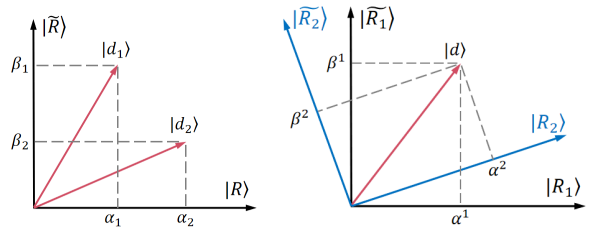
In summary, our contributions are as follows:

- We apply, for the first time to the best of our knowledge, the quantum theory to RA problem. We present a quantitative interpretation of the generation of basic rankings, and analyze the theoretical motivation for the classic linear weighting idea in previous RA methods.
- We propose a QI-IRA method, which introduces a new interactive fusion weight updating scheme inspired by quantum observation theory. Since QI-IRA is extremely light-weighted and parameter-efficient, it can therefore be easily deployed to a variety of application scenarios.
- We conduct extensive comparative experiments on four image and two video re-ID datasets, on which QI-IRA achieves consistent superiority over unsupervised, fully-supervised and interactive RA baselines. Particularly, QI-IRA can outperform fully-supervised RA methods with as low as only two feedbacks per query on Market1501, MARS and DukeMTMC-VideoReID datasets.

Related Work

Depending on the availability of labeled data (Oliveira et al. 2020), the mainstream RA research can be generally classified as unsupervised and fully-supervised methods. Unsupervised methods have no access to labeled data and rely only on heuristic metrics, such as mean, median, and etc., to aggregate basic rankings. The Comb* family (Fox and Shaw 1994) computes positional scores of items and combines them using various strategies. ER is based on the half-quadratic (HQ) theory and determines an optimal weight for each methods. CG (Xiao et al. 2021) develops a graph-based method and sorts items based on the ratio of out- and in-degrees (ROID). Due to lack of external supervision, unsupervised methods cannot directly optimize the final RA performance (Nasteski 2017).

Fully-supervised methods, in contrast, utilize the labeled data to improve the accuracy. As Learning-To-Rank (LTR) methods, RankNet (Burges et al. 2005) is a pairwise method, LambdaRank (Burges, Ragno, and Le 2006) introduces Lambda gradients based on it, and LambdaMART (Wu et al. 2010) combines Lambda gradients and MART. CSRA evaluates the reliability score of each ranking and aggregates them accordingly. However, fully-supervised methods require massive labeling information that is usually expen-



(a) Ranking by one method (b) Ranking by two methods

Figure 3: Generation of multiple basic rankings.

sive or simply or unavailable in the practical applications (Dourado, Pedronette, and da Silva Torres 2019).

To resolve the dilemma between performance and cost faced by unsupervised and fully-supervised RA methods, IRA is proposed to fuse basic rankings via linear weighting guided by relevance feedback, which can improve the performance under a small amount of labeling effort. However, IRA does not properly analyze the rationality of the linear weighting idea, and its weight updating rule heavily relies on empirically set parameters, which are sub-optimal in the comparison with QI-IRA (Table 2).

Quantum Preliminaries

Quantum probability formalism is based on vector spaces using Dirac Notations, known as "bra-ket" notation. The state vector of a microscopic particle is denoted by a unit vector $|d\rangle$ (called as ket), and its conjugate transpose is denoted as $\langle d|$ (called as bra). According to quantum observation theory, any observation operation corresponds to an hermitian operator whose eigenvectors form a set of orthogonal bases $\{|R_m\rangle\}$ in the state space. The state observation of a microscopic particle will cause the particle to collapse from its original state to one of the eigenstates of the observation operator with a probability. In turn, the state vector can be expressed as the linear superposition of the base vectors: $|d\rangle = \sum_{m=1} c_m |R_m\rangle$, where c_m is the probability amplitude and $\sum_{m=1} |c_m|^2 = 1$. The probability amplitude is defined as $c_m = \langle R_m|d\rangle$, which is the projection of $|d\rangle$ onto $|R_m\rangle$, and its corresponding collapse probability is $p(\langle R_m|d\rangle) = |c_m|^2 = \|\langle R_m|d\rangle\|^2$. There are different physical quantities of a microscopic particle that can be observed, which means that the same particle state can be represented and observed under different orthogonal bases.

Method

This section discusses the technical details of QI-IRA. We first present a quantitative interpretation of the generation of basic rankings, and then analyze the quantum-inspired aggregation method for basic rankings.

Generation of Basic Rankings

Inspired by the quantum observation theory, in re-ID problem we consider a gallery image as a microscopic particle, whose state can be represented as a vector in the retrieval space. And a basic re-ID method corresponds to a quantum observation operation, which has a set of orthogonal bases. Since re-ID is intrinsically an information retrieval problem, hence the state space can be simplified by two type of bases, one is relevant base $|R\rangle$ and the other is irrelevant base $|\tilde{R}\rangle$. Figure 3a shows the ranking (observation) of two gallery images (microscopic particles) under the same re-ID method (observation operator), from which we care about the collapsed projections of two state vectors $|d_1\rangle$ and $|d_2\rangle$ to the relevant base $|R\rangle$. Figure 3b shows the rankings (observations) of the same gallery image (microscopic particle) under two re-ID methods (observation operators), which is essentially the collapsed projections of the same state vector $|d\rangle$ to two relevant bases $|R_1\rangle$ and $|R_2\rangle$. In this case, the algebraic representation of the state vector of the gallery image can be formulated as:

$$\begin{aligned} |d\rangle &= \alpha^1|R_1\rangle + \beta^1|\widetilde{R_1}\rangle \\ &= \alpha^2|R_2\rangle + \beta^2|\widetilde{R_2}\rangle \end{aligned} \tag{1}$$

where $\{\alpha^m\}$ and $\{\beta^m\}$ respectively correspond to the probability amplitude of the state of the gallery image under the m-th observation's bases $\{|R_m\rangle\}$ and $\{|\widetilde{R_m}\rangle\}$. In this paper, the probability amplitude is defined as $\alpha^m = \sqrt{s^m}$, where s^m is the ranking score of the gallery image under the m-th re-ID method. In turn, $s^m = |\alpha^m|^2 = p(\langle R_m|d\rangle)$, indicating that the ranking score s^m is the collapse probability of the gallery image under the m-th re-ID method.

Since the same gallery image can be described by bases of various observation operators (re-ID methods), corresponding to different collapse probabilities (ranking scores), thus generating multiple basic rankings.

Aggregation of Basic Rankings

Assuming that the state of a gallery image is $|d\rangle$, and the relevant base of a re-ID method is $|R\rangle$, its ranking score (collapse probability) can be reformulated with Gleason's theorem (Van Rijsbergen 2004):

$$\begin{aligned} s &= \|\langle R|d\rangle\|^2 \\ &= \mathrm{tr}[(|d\rangle\langle d|) \cdot (|R\rangle\langle R|)] \\ &= \mathrm{tr}[D\rho] \end{aligned} \tag{2}$$

where $\rho = |R\rangle\langle R|$ is a basic density operator corresponding to a single re-ID method, and $D = |d\rangle\langle d|$ is the state operator corresponding to the gallery image. According to the previous discussion, a single re-ID method represents an incomplete observation. In order to obtain more accurate ranking results, we adopt linear weighting to recover the latent complete observation, whose state can be formulated as:

$$\rho = \sum_{m=1} w_m |R_m\rangle\langle R_m| \tag{3}$$

where w_m is the fusion weights to allocate the individual importance of incomplete observations in the complete one. Substituting Eq. 3 into Eq. 2, the aggregated ranking score s^* of the gallery image under the complete observation can

Algorithm 1: QI-IRA

Input: Query q and basic ranking scores $\mathbb{S} = \{\mathbf{s}^m\}_{m=1}^M$
Output: Aggregated ranking scores $\mathbf{s}*$

1: Adopt Mean RA method to generate initial ranking
2: **while** not a satisfied ranking **do**
3: Collect Ψ_+^q and Ψ_-^q using relevance feedback
4: Calculate new fusion weights $\{\bar{w}_m^t\}_{m=1}^M$ by Eq. 5
5: Update the fusion weights $\{w_m^t\}_{m=1}^M$ by Eq. 6
6: **end while**
7: Compute aggregated ranking scores $\mathbf{s}^*$ by Eq. 4

be expressed as:

$$\begin{aligned} s^* &= \sum_{m=1} w_m \mathrm{tr}[D|R_m\rangle\langle R_m|] \\ &= \sum_{m=1} w_m \| \langle R_m|d\rangle \|^2 \\ &= \sum_{m=1} w_m s^m \end{aligned} \tag{4}$$

where $s^m = \| \langle R_m|d\rangle \|^2$ is the ranking score of the gallery image under the m-th re-ID method. This results motivates a classic linear weighting idea for QI-IRA. But, in the following, we will find that QI-IRA adopts a simple but effective scheme to update the fusion weights empowered by quantum observation theory.

In order to calculate the fusion weights, we adopts a relevant feedback strategy similar to IRA. The basic idea is introducing relevance feedback, *i.e.*, user judgement, to evaluate the current ranking result. From this we can increase the weights of basic rankings that are consistent with the current feedback set, and meanwhile, decrease the weights of basic rankings that are inconsistent with it. However, unlike IRA, QI-IRA does not use the standard deviation of ranking scores to update the fusion weights $\{w_m\}$, but rather updates it based on the collapse probabilities of the feedback samples on basic re-ID methods, *i.e.*, basic ranking scores. Specifically, the fusion weight of the m-th basic ranking in the t-th round is:

$$\begin{aligned} \bar{w}_m^t &= \frac{1}{|\Psi_+^q|} \sum_{d_i \in \Psi_+^q} \| \langle R_m|d_i\rangle \|^2 - \frac{1}{|\Psi_-^q|} \sum_{d_j \in \Psi_-^q} \| \langle R_m|d_j\rangle \|^2 \\ &= \frac{1}{|\Psi_+^q|} \sum_{d_i \in \Psi_+^q} |\alpha_i^m|^2 - \frac{1}{|\Psi_-^q|} \sum_{d_j \in \Psi_-^q} |\alpha_j^m|^2 \\ &= \frac{1}{|\Psi_+^q|} \sum_{d_i \in \Psi_+^q} s_i^m - \frac{1}{|\Psi_-^q|} \sum_{d_j \in \Psi_-^q} s_j^m \end{aligned} \tag{5}$$

where $|\Psi_+^q|$ and $|\Psi_-^q|$ are the number of gallery images in the positive and negative feedback sets, respectively.

In the very beginning, since there is no user feedback, we adopt the Mean RA method (Burges et al. 2011) by default to obtain the initial ranking. Then, in the following iterations, the user can give feedback on the top-K items in the current round[1], and update the fusion weight using with a momentum as:

$$w_m^t = \gamma \times \bar{w}_m^t + (1-\gamma) \times w_m^{t-1} \tag{6}$$

where γ is the only model parameter in QI-IRA, which controls the degree of weight updating. We will investigate its dynamic performance in the experimental section. The whole process is shown in Algorithm 1.

Before we start the experimental section, we need to emphasize again that although the linear weighting idea has been widely used in previous RA methods like ER, CSRA and IRA, none of them have theoretically analyzed or discussed the rationality of this idea. Moreover, some methods heavily rely on empirically set model parameters in the process of learning weights, which may be sub-optimal on new datasets. In contrast, QI-IRA provides a new interpretation of the generation and aggregation of basic rankings with the help of quantum theory. Not only its weight updating mechanism is simple (Algorithm 1) and efficient (Table 3), but also the only model parameter γ is insensitive to the final RA performance (Figure 5), which enables QI-IRA to achieve better and more robust performance than the existing RA methods on different datasets (Table 1,Table 2 and Figure 4).

Experiments

Dataset and Metrics

We evaluate our method on four image re-ID datasets (Market1501 (Zheng et al. 2015), DukeMTMC-reID (Ristani et al. 2016) and CUHK03 detected and labeled (Li et al. 2014)), and two video re-ID datasets (MARS (Zheng et al. 2016) and DukeMTMC-VideoReID (Wu et al. 2018)), which are all popular benchmarks evaluated in various person re-ID studies. We use official training sets to train basic re-ID and fully-supervised RA methods, and use official test sets to evaluate all RA methods. Two classic performance evaluation criteria, *i.e.*, Rank@1 and mAP (Ye et al. 2021) are used in the following experiments.

Implementation Details

We use QI-IRA(K, T) to represent our method, where K denotes the number of feedback samples in each interaction round and T denotes the maximum number of interaction rounds. So $K \times T$ is the total number of feedback samples. In the weight adjustment step of QI-IRA, the fusion factor γ is set to 0.7 for the weight of the new round when fusing the weights calculated from two adjacent interaction rounds.

For fully-supervised methods, we use the original codes directly in our experiments (LambdaMART[2], LambdaRank and RankNet[3]).

Comparison with Unsupervised RA Methods

We compare QI-IRA with sixteen representative unsupervised RA methods, including the Comb* family (CombMIN, CombMAX, CombSUM, CombANZ, CombMNZ),

[1] If some samples in the top-K items have been labeled in the current round, it will be postponed until K new samples are labeled

[2] https://github.com/saleed/lamdamart

[3] https://github.com/houchenyu/L2R/tree/master

		Image Dataset								Video Dataset			
		Market1501		DukeMTMC-reID		CUHK03 (detected)		CUHK03 (labeled)		MARS		DukeMTMC-VideoReID	
Method	**Venue**	**R@1**	**mAP**	**R@1**	**mAP**	**R@1**	**mAP**	**R@1**	**mAP**	**R@1**	**mAP**	**R@1**	**mAP**
CombMIN	NIST SP'94	96.08	94.26	91.83	**88.88**	79.86	80.33	81.71	82.62	91.01	85.42	95.44	94.41
CombMAX	NIST SP'94	95.64	91.29	89.45	82.62	81.29	78.79	83.79	81.68	90.20	84.67	95.87	94.59
CombSUM	NIST SP'94	96.67	92.88	92.06	85.68	83.00	81.06	86.36	84.35	91.21	85.97	95.44	94.76
CombANZ	NIST SP'94	94.33	86.40	85.77	76.74	75.50	74.80	81.21	79.21	87.12	78.79	95.58	94.22
CombMNZ	NIST SP'94	96.67	89.14	92.15	81.66	83.07	81.78	86.36	84.87	91.16	82.77	95.44	94.77
BordaCount	SIGIR'01	96.67	92.88	92.06	85.68	83.00	81.06	86.36	84.35	91.21	85.97	95.44	94.76
Dowdall	IPSR'02	96.65	93.50	91.20	86.79	84.50	82.30	**87.21**	85.13	91.21	86.28	95.87	95.22
Median	SIGMOD'03	96.62	93.37	91.92	85.93	**85.29**	82.84	87.07	85.09	91.36	86.58	**96.58**	95.78
RRF	SIGIR'09	**96.88**	93.41	**92.46**	86.74	83.64	81.89	86.64	84.85	**91.41**	86.46	95.44	94.89
iRANK	JIST'10	96.50	94.35	92.32	88.17	84.86	83.85	87.00	86.38	91.26	86.87	**96.58**	**95.81**
Mean	PMLR'11	96.44	94.55	92.24	88.57	84.71	84.40	86.93	86.91	91.26	87.15	**96.58**	**95.81**
HPA	ECIR'20	96.44	**94.74**	92.10	**89.09**	81.64	82.23	83.57	84.52	91.26	86.59	96.30	95.52
PostNDCG	ECIR'20	96.47	92.09	91.47	84.47	82.50	79.72	83.71	81.92	90.05	83.71	96.01	95.31
ER	OMEGA'20	96.64	92.89	92.15	85.68	83.14	81.09	86.36	84.42	91.26	85.98	96.15	95.15
CG	JORS'21	96.67	92.88	92.06	85.68	83.00	81.06	86.36	84.35	91.21	85.97	95.44	94.76
DIBRA	LSA'22	96.44	**94.58**	92.37	88.69	84.64	84.39	86.86	86.91	91.26	**87.19**	**96.58**	95.77
QI-IRA(1,1)	Ours	96.44	94.55	92.37	88.63	84.79	**84.41**	87.07	**86.93**	91.26	87.15	**96.58**	**95.81**
QI-IRA(2,1)	Ours	**97.15**	**94.74**	**92.68**	88.74	**84.93**	**84.54**	**87.57**	**87.26**	**93.37**	**87.72**	**98.52**	**97.01**

Table 1: Comparison results between QI-IRA(K, T) and unsupervised methods. Here the bold underline indicates the best value and the bold indicates the second best.

	Method	Image Dataset								Video Dataset			
		Market1501		DukeMTMC-reID		CUHK03 (detected)		CUHK03 (labeled)		MARS		DukeMTMC-VideoreID	
		T=1	T=3	T=1	T=3	T=1	T=3	T=1	T=3	T=1	T=3	T=1	T=3
K=1	$\mathrm{IRA_R}$(K,T)	92.85	93.02	86.75	87.03	82.62	83.11	85.53	86.04	85.43	86.24	93.46	93.60
	$\mathrm{IRA_S}$(K,T)	**94.56**	94.76	88.58	88.56	**84.42**	84.36	86.91	87.29	**87.15**	87.92	**95.81**	96.17
	QI-IRA(K,T)	94.55	**95.04**	**88.63**	**88.75**	84.41	**84.87**	**86.93**	**87.37**	**87.15**	**88.31**	**95.81**	**97.82**
K=3	$\mathrm{IRA_R}$(K,T)	94.70	95.01	88.82	89.40	84.82	89.64	87.19	91.41	87.90	90.13	96.04	97.05
	$\mathrm{IRA_S}$(K,T)	94.77	94.95	88.67	88.91	84.48	89.15	87.48	91.41	87.98	90.07	95.88	96.81
	QI-IRA(K,T)	**94.98**	**96.10**	**88.91**	**90.21**	**84.87**	**92.26**	**87.55**	**93.89**	**88.30**	**91.68**	**97.81**	**99.18**
K=5	$\mathrm{IRA_R}$(K,T)	94.83	95.28	88.90	90.07	86.08	93.12	88.03	93.89	88.67	91.39	96.65	97.21
	$\mathrm{IRA_S}$(K,T)	94.77	95.09	88.85	89.66	86.06	92.57	88.82	94.30	88.86	91.41	96.35	96.91
	QI-IRA(K,T)	**95.35**	**96.92**	**89.24**	**91.85**	**87.11**	**95.55**	**89.40**	**96.46**	**89.70**	**93.42**	**98.73**	**99.36**

Table 2: Comparison results of mAP (%) between QI-IRA and IRA. Here the bold indicates the best value.

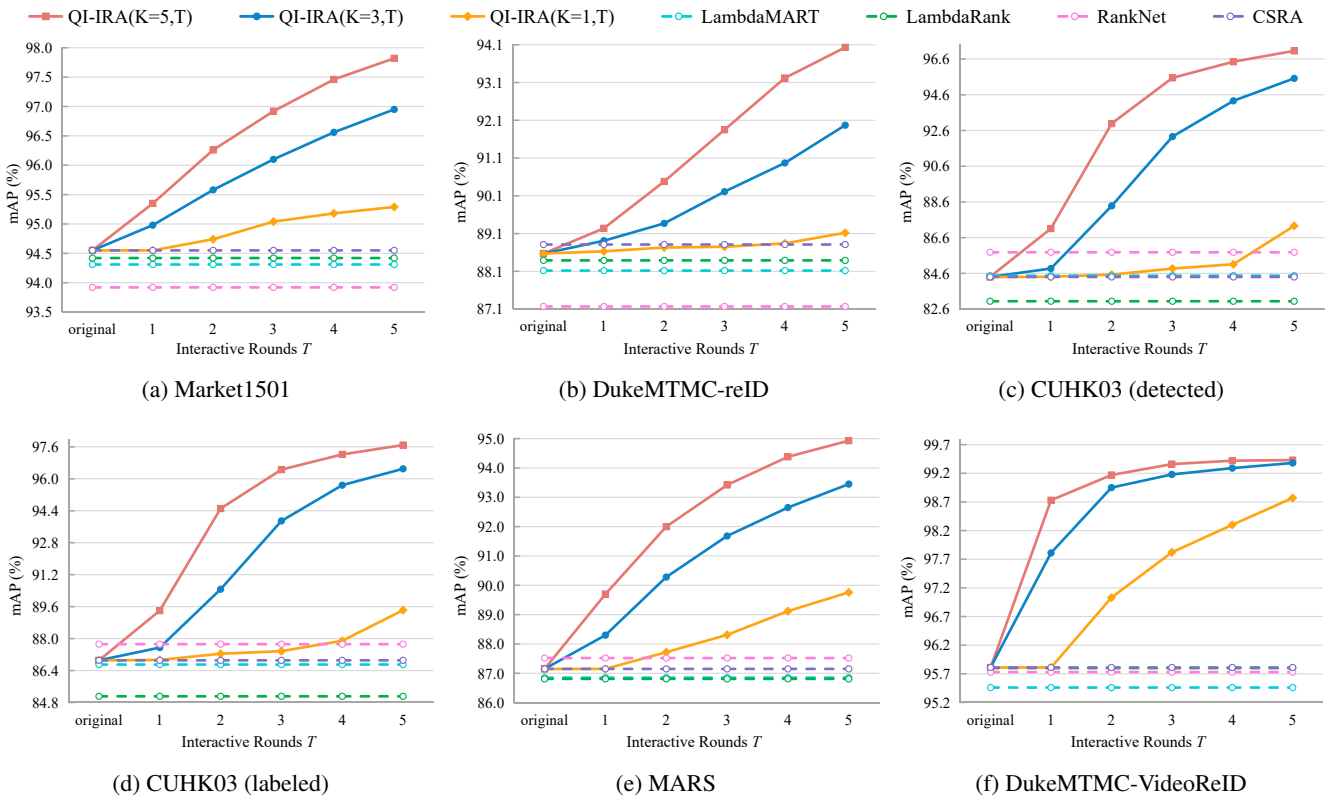
BordaCount (Aslam and Montague 2001), Dowdall (Reilly 2002), Median (Fagin, Kumar, and Sivakumar 2003), RRF (Cormack, Clarke, and Buettcher 2009), iRANK (Wei, Li, and Liu 2010), Mean (Burges et al. 2011), HPA (Fujita, Kobayashi, and Okumura 2020), PostNDCG (Fujita, Kobayashi, and Okumura 2020), ER, CG and DIBRA (Akritidis et al. 2022). The results on six re-ID datasets are shown in Table 1.

In the Table 1, at K = 1, QI-IRA performs best in mAP criterion on three of six datasets. At K = 2, QI-IRA ranks the first position in the majority of cases. The possible reason why QI-IRA is not the best in mAP on DukeMTMC-reID is that there are more items in the basic rankings on DukeMTMC-reID, which limits the performance improvement when $K \times T$ is small. Still, on this dataset, QI-IRA ranks the first in Rank@1 and the third in mAP at K = 2.

Comparison with Interactive RA Method

We compare QI-IRA with the interactive method IRA. IRA uses the positive feedback information from interaction to adjust weights, and its strategy consists of two types: ranking position based ($\mathrm{IRA_R}$) and score standard deviation based ($\mathrm{IRA_S}$). To ensure fairness, we test multiple combinations of various feedback samples and interaction rounds.

Table 2 shows the comparative results of mAP between QI-IRA and IRA on six re-ID datasets. It can be seen that: (1) QI-IRA outperforms $\mathrm{IRA_R}$ on all datasets under the same experimental setup. In particular, at K = 5 and T = 3, the mAP of both methods has been better than 96.90% on DukeMTMC-VideoReID, where QI-IRA has mAP = 99.36%, still an improvement of 2.15% over that of $\mathrm{IRA_R}$. (2) In the vast majority of cases, the mAP of QI-IRA is higher than that of $\mathrm{IRA_S}$. When K = 5 and T = 3, the mAP of QI-IRA is 2.98% higher than that of $\mathrm{IRA_S}$ on CUHK03



(a) Market1501 (b) DukeMTMC-reID (c) CUHK03 (detected)

(d) CUHK03 (labeled) (e) MARS (f) DukeMTMC-VideoReID

Figure 4: Comparison results between QI-IRA and fully-supervised methods.

Method	mAP	Train (min)	Test		Avg. (s)
			Int. (s)	Agg. (ms)	
CSRA	84.40	209.63	0.00	3.24	8.98
LambdaMART	84.47	190.08	0.00	3.02	8.15
LambdaRank	83.04	162.06	0.00	1.49	6.95
RankNet	85.78	148.30	0.00	6.08	6.36
$\text{IRA}_{\text{R}}(1,1)$	82.62	0.00	3.00	1.14	3.00
$\text{IRA}_{\text{S}}(1,1)$	84.42	0.00	3.00	0.79	3.00
QI-IRA(1,1)	84.41	0.00	3.00	0.71	3.00

Table 3: Average time cost of QI-IRA, IRA and fully-supervised methods on CUHK03 (detected).

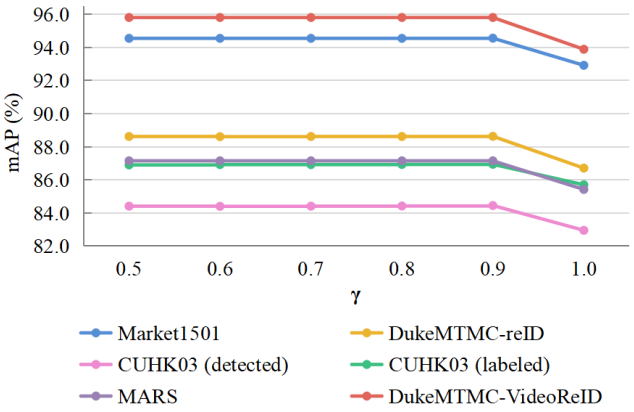


Figure 5: Sensitivity curves for parameter γ at $K = T = 1$.

(detected). (3) QI-IRA demonstrates the superiority and robustness in varying datasets. The performance of IRA is constrained due to heuristic rules for fusion weights' update. On the contrary, QI-IRA puts forward a new and robust strategy guided by quantum theory, making more effective use of the small amount of feedback information.

Comparison with Fully-Supervised RA Methods

We conduct comparison experiments for QI-IRA and four fully-supervised RA methods (CSRA, LambdaMART, LambdaRank and RankNet).

Figure 4 shows mAP results of QI-IRA and fully-supervised methods. On six re-ID datasets, the performance of QI-IRA initial rankings doesn't exceed that of the best fully-supervised method. QI-IRA outperforms all fully-supervised methods at $K = 1$ with only two rounds of interaction on Market1501, MARS, and DukeMTMC-VideoReID, with four rounds of interaction on DukeMTMC-reID and CUHK03 (labeled), and with five rounds of interaction on CUHK03 (detected). Moreover, as the number of feedback samples per round or interaction rounds increases, the performance of QI-IRA grows significantly on these datasets.

Figure 6: Visualized examples of the Top-10 aggregated results at $K = 1$ on the DukeMTMC-reID dataset. Images with green bounding boxes indicate galleries that match the query.

Comparison on the Time Cost

Low time cost is a unique advantage of interactive methods over fully-supervised methods. For fully-supervised methods, time costs include training time and aggregating time. For QI-IRA and IRA, time costs consist of interacting time and aggregating time. Estimating a 3-second average interacting time per gallery, the total interacting time can be calculated by $K \times T$. Because the performance of QI-IRA varies discretely with $K \times T$, we compare the time cost of methods when their performance is similar in mAP.

Table 3 shows the average time cost per query of QI-IRA, IRA and fully-supervised methods on CUHK03 (detected). It's obvious that the average time cost required for QI-IRA is lower than that of fully-supervised methods. Compared with them, QI-IRA can start RA process and obtain results more quickly in practical applications. In addition, compared with IRA, QI-IRA approximates or even exceeds the performance of IRA with less time cost, reflecting that QI-IRA can utilize feedback information more effectively.

Parameter Sensitivity Analysis

The parameter γ is the fusion factor for the weight of the new round when fusing the weights calculated from two adjacent interaction rounds. Figure 5 shows its sensitivity curves on six re-ID datasets at $K \times T = 1$ as an example. The γ curves remain stable between 0.5 and 0.9, indicating little impact on aggregation performance. Overall, to be on the safe side, we set γ to 0.7 as the default in QI-IRA implementation.

Qualitative Result

Several visualized examples of changes in Top-10 samples between the initial ranking and the aggregated ranking at $K = 1$ on the DukeMTMC-reID dataset after one round of feedback are shown in Figure 6. After one interaction, more Top-10 samples match the query, fewer samples remain unmatched, and the rank positions of unmatched samples shift back significantly. This illustrates the effectiveness of fusion weight adjustment strategy of QI-IRA.

Conclusion

Inspired by quantum theory, we propose an interactive ranking aggregation (QI-IRA) method, which addresses the limitations of existing RA methods. Our method interprets and models the generation and aggregation of multiple basic rankings, presents a robust strategy for learning fusion weights through interaction. Extensive comparative experiments show the consistent superiority of QI-IRA over existing RA methods. RA is a fundamental technology used in many fields. In the future we will apply QI-IRA to more applications through research and improvement.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (No. U1903214, 62372339, 61876135, 61976163), the Ministry of Education Industry-University Cooperative Education Project (No. 202102246004, 220800006041043), the Key R&D Program of Hubei Province (No. 2023BAB077), and the Fundamental Research Funds for the Central Universities (No. 2042023kf0219). This work was supported by Ant Group through CCF-Ant Research Fund. The numerical calculations in this paper have been done on the supercomputing system in the Supercomputing Center of Wuhan University.

References

Akritidis, L.; Fevgas, A.; Bozanis, P.; and Manolopoulos, Y. 2022. An unsupervised distance-based model for weighted rank aggregation with list pruning. *Expert Systems with Applications*, 202: 117435.

Aslam, J. A.; and Montague, M. 2001. Models for metasearch. In *Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval*, 276–284.

Bałchanowski, M.; and Boryczka, U. 2022. Collaborative Rank Aggregation in Recommendation Systems. *Procedia Computer Science*, 207: 2213–2222.

Borda, J.-C. d. 1781. Mémoire sur les élections au scrutin: Histoire de l'Académie Royale des Sciences. *Paris, France*, 12.

Burges, C.; Ragno, R.; and Le, Q. 2006. Learning to rank with nonsmooth cost functions. *Advances in neural information processing systems*, 19.

Burges, C.; Shaked, T.; Renshaw, E.; Lazier, A.; Deeds, M.; Hamilton, N.; and Hullender, G. 2005. Learning to rank using gradient descent. In *Proceedings of the 22nd international conference on Machine learning*, 89–96.

Burges, C.; Svore, K.; Bennett, P.; Pastusiak, A.; and Wu, Q. 2011. Learning to rank using an ensemble of lambda-gradient models. In *Proceedings of the learning to rank Challenge*, 25–35. PMLR.

Cho, Y.; Kim, W. J.; Hong, S.; and Yoon, S.-E. 2022. Part-based pseudo label refinement for unsupervised person re-identification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 7308–7318.

Cormack, G. V.; Clarke, C. L.; and Buettcher, S. 2009. Reciprocal rank fusion outperforms condorcet and individual rank learning methods. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, 758–759.

Dai, Z.; Chen, M.; Gu, X.; Zhu, S.; and Tan, P. 2019. Batch dropblock network for person re-identification and beyond. In *Proceedings of the IEEE/CVF international conference on computer vision*, 3691–3701.

Deng, K.; Han, S.; Li, K. J.; and Liu, J. S. 2014. Bayesian aggregation of order-based rank data. *Journal of the American Statistical Association*, 109(507): 1023–1039.

Dourado, I. C.; Pedronette, D. C. G.; and da Silva Torres, R. 2019. Unsupervised graph-based rank aggregation for improved retrieval. *Information Processing & Management*, 56(4): 1260–1279.

Fagin, R.; Kumar, R.; and Sivakumar, D. 2003. Efficient similarity search and classification via rank aggregation. In *Proceedings of the 2003 ACM SIGMOD international conference on Management of data*, 301–312.

Fox, E. A.; and Shaw, J. A. 1994. Combination of multiple searches. *NIST special publication SP*, 243.

Fu, D.; Chen, D.; Bao, J.; Yang, H.; Yuan, L.; Zhang, L.; Li, H.; and Chen, D. 2021. Unsupervised pre-training for person re-identification. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 14750–14759.

Fujita, S.; Kobayashi, H.; and Okumura, M. 2020. Unsupervised Ensemble of Ranking Models for News Comments Using Pseudo Answers. In *Advances in Information Retrieval: 42nd European Conference on IR Research, ECIR 2020, Lisbon, Portugal, April 14–17, 2020, Proceedings, Part II 42*, 133–140. Springer.

Ge, Y.; Zhu, F.; Chen, D.; Zhao, R.; et al. 2020. Self-paced contrastive learning with hybrid memory for domain adaptive object re-id. *Advances in Neural Information Processing Systems*, 33: 11309–11321.

Herzog, F.; Ji, X.; Teepe, T.; Hörmann, S.; Gilg, J.; and Rigoll, G. 2021. Lightweight multi-branch network for person re-identification. In *2021 IEEE International Conference on Image Processing (ICIP)*, 1129–1133. IEEE.

Huang, J.; Liang, C.; Zhang, Y.; Wang, Z.; and Zhang, C. 2022. Ranking Aggregation with Interactive Feedback for Collaborative Person Re-identification.

Li, W.; Zhao, R.; Xiao, T.; and Wang, X. 2014. Deepreid: Deep filter pairing neural network for person re-identification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 152–159.

Liu, J.; Sun, Y.; Zhu, F.; Pei, H.; Yang, Y.; and Li, W. 2022. Learning memory-augmented unidirectional metrics for cross-modality person re-identification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 19366–19375.

Luo, H.; Gu, Y.; Liao, X.; Lai, S.; and Jiang, W. 2019. Bag of tricks and a strong baseline for deep person re-identification. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, 0–0.

Luo, X.; Ouyang, Z.; Du, N.; Song, J.; and Wei, Q. 2021. Cross-Domain Person Re-Identification Based on Feature Fusion. *IEEE Access*, 9: 98327–98336.

Mansouri, N.; Ammar, S.; and Kessentini, Y. 2021. Re-ranking person re-identification using attributes learning. *Neural Computing and Applications*, 33: 12827–12843.

Mohammadi, M.; and Rezaei, J. 2020. Ensemble ranking: Aggregation of rankings produced by different multi-criteria decision-making methods. *Omega*, 96: 102254.

Nasteski, V. 2017. An overview of the supervised machine learning methods. *Horizons. b*, 4: 51–62.

Oliveira, S. E.; Diniz, V.; Lacerda, A.; Merschmanm, L.; and Pappa, G. L. 2020. Is rank aggregation effective in recommender systems? an experimental analysis. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 11(2): 1–26.

Quispe, R.; and Pedrini, H. 2021. Top-db-net: Top dropblock for activation enhancement in person re-identification. In *2020 25th International conference on pattern recognition (ICPR)*, 2980–2987. IEEE.

Reilly, B. 2002. Social choice in the south seas: Electoral innovation and the borda count in the pacific island countries. *International Political Science Review*, 23(4): 355–372.

Ristani, E.; Solera, F.; Zou, R.; Cucchiara, R.; and Tomasi, C. 2016. Performance measures and a data set for multi-target, multi-camera tracking. In *Computer Vision–ECCV*

2016 Workshops: Amsterdam, The Netherlands, October 8-10 and 15-16, 2016, Proceedings, Part II, 17–35. Springer.

Van Rijsbergen, C. J. 2004. *The geometry of information retrieval*. Cambridge University Press.

Wang, B.; Law, A.; Regan, T.; Parkinson, N.; Cole, J.; Russell, C. D.; Dockrell, D. H.; Gutmann, M. U.; and Baillie, J. K. 2022. Systematic comparison of ranking aggregation methods for gene lists in experimental results. *Bioinformatics*, 38(21): 4927–4933.

Wang, H.; Wang, G.; Li, Y.; Zhang, D.; and Lin, L. 2020. Transferable, controllable, and inconspicuous adversarial attacks on person re-identification with deep mis-ranking. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 342–351.

Wei, F.; Li, W.; and Liu, S. 2010. iRANK: A rank-learn-combine framework for unsupervised ensemble ranking. *Journal of the American Society for Information Science and Technology*, 61(6): 1232–1243.

Wu, Q.; Burges, C. J.; Svore, K. M.; and Gao, J. 2010. Adapting boosting for information retrieval measures. *Information Retrieval*, 13: 254–270.

Wu, S.; Huang, C.; Li, L.; and Crestani, F. 2019. Fusion-based methods for result diversification in web search. *Information Fusion*, 45: 16–26.

Wu, Y.; Lin, Y.; Dong, X.; Yan, Y.; Ouyang, W.; and Yang, Y. 2018. Exploit the unknown gradually: One-shot video-based person re-identification by stepwise learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 5177–5186.

Xiao, Y.; Deng, H.-Z.; Lu, X.; and Wu, J. 2021. Graph-based rank aggregation method for high-dimensional and partial rankings. *Journal of the Operational Research Society*, 72(1): 227–236.

Yanofsky, N. S.; and Mannucci, M. A. 2008. *Quantum computing for computer scientists*. Cambridge University Press.

Ye, M.; Liang, C.; Yu, Y.; Wang, Z.; Leng, Q.; Xiao, C.; Chen, J.; and Hu, R. 2016. Person reidentification via ranking aggregation of similarity pulling and dissimilarity pushing. *IEEE Transactions on Multimedia*, 18(12): 2553–2566.

Ye, M.; Shen, J.; Lin, G.; Xiang, T.; Shao, L.; and Hoi, S. C. 2021. Deep learning for person re-identification: A survey and outlook. *IEEE transactions on pattern analysis and machine intelligence*, 44(6): 2872–2893.

Yu, Y.; Liang, C.; Ruan, W.; and Jiang, L. 2020. Crowdsourcing-Based Ranking Aggregation for Person Re-Identification. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1933–1937. IEEE.

Yuan, Y.; Chen, W.; Yang, Y.; and Wang, Z. 2020. In defense of the triplet loss again: Learning robust person re-identification with fast approximated triplet loss and label distillation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 354–355.

Zahra, A.; Perwaiz, N.; Shahzad, M.; and Fraz, M. M. 2023. Person re-identification: A retrospective on domain specific open challenges and future trends. *Pattern Recognition*, 109669.

Zhang, S.; Yin, Z.; Wu, X.; Wang, K.; Zhou, Q.; and Kang, B. 2021. FPB: feature pyramid branch for person re-identification. *arXiv preprint arXiv:2108.01901*.

Zhang, Y.; Qian, Q.; Liu, C.; Chen, W.; Wang, F.; Li, H.; and Jin, R. 2022. Graph convolution for re-ranking in person re-identification. In *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2704–2708. IEEE.

Zheng, L.; Bie, Z.; Sun, Y.; Wang, J.; Su, C.; Wang, S.; and Tian, Q. 2016. Mars: A video benchmark for large-scale person re-identification. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part VI 14*, 868–884. Springer.

Zheng, L.; Shen, L.; Tian, L.; Wang, S.; Wang, J.; and Tian, Q. 2015. Scalable person re-identification: A benchmark. In *Proceedings of the IEEE international conference on computer vision*, 1116–1124.