Re-ranking vehicle re-identification with orientation-guide query expansion

Xue Zhang\textsuperscript{1}, Xiushan Nie\textsuperscript{1}, Ziruo Sun\textsuperscript{2}, Xiaofeng Li\textsuperscript{1}, Chuntao Wang\textsuperscript{3}, Peng Tao\textsuperscript{3} and Sumaira Hussain\textsuperscript{4}

Abstract
Vehicle re-identification, which aims to retrieve information regarding a vehicle from different cameras with non-overlapping views, has recently attracted extensive attention in the field of computer vision owing to the development of smart cities. This task can be regarded as a type of retrieval problem, where re-ranking is important for performance enhancement. In the vehicle re-identification ranking list, images whose orientations are dissimilar to that of the query image must preferably be optimized on priority. However, traditional methods are incompatible with such samples, resulting in unsatisfactory vehicle re-identification performances. Therefore, in this study, we propose a vehicle re-identification re-ranking method with orientation-guide query expansion to optimize the initial ranking list obtained by a re-identification model. In the proposed method, we first find the nearest neighbor image whose orientation is dissimilar to the queried image and then fuse the features of the query and neighbor images to obtain new features for information retrieval. Experiments are performed on two public data sets, VeRi-776 and VehicleID, and the effectiveness of the proposed method is confirmed.

Keywords
Vehicle re-identification, re-ranking, orientation-guide, feature fusion, query expansion

Date received: 21 September 2021; accepted: 31 October 2021
Handling Editor: Yanjiao Chen

Introduction
Owing to the increasing development of video surveillance and public security methods and systems, there is a growing need for vehicle re-identification (re-ID) from images. However, this is a challenging task in computer vision and can be regarded as an information retrieval problem. Given a query vehicle image, a vehicle re-ID method attempts to find all images containing that vehicle across multiple non-overlapping cameras. It can be seen from the initial sensor-based methods\textsuperscript{1–3} for re-ID, hand-crafted-feature-based methods,\textsuperscript{4–9} and deep-feature-based methods\textsuperscript{10–15} that the ability to express the acquired features from images is rapidly improving. However, owing to the range of camera capture angles, the orientations of the vehicle images may differ, and these vehicle images often have differences in visual effects. As shown in Figure 1, images whose orientations are different from that of the query images may rank lower than those whose orientations are similar to that of the query image. To address this problem of orientation variation, in certain previous studies, components were primarily added in the training phase to learn the embedded features.
robustly, thereby increasing the training complexity. Thus, in this study, we propose a simple but effective method to tackle the issue in the post-processing stage.

In the post-processing stage, re-ranking is used as an effective method to optimize the retrieval ranking. Average query expansion (AQE)\(^{16}\) is a type of re-ranking method based on the \(k\)-nearest neighbor principle, where the new query features are constructed by averaging the old features and the top-\(k\) features in the returned list are ranked in order before retrieval is performed again to obtain the final result. Jegou et al.\(^{17}\) used the corresponding neighborhood and proposed the contextual dissimilarity measure (CDM) to improve the re-ID performance. Leng et al.\(^{18}\) first proposed the relative information of the nearest neighborhood of each image to improve re-ranking. Sparse contextual activation (SCA)\(^{19}\) simply completed re-ranking tasks through vectors comparison under the generalized Jaccard metric. Zhong et al.\(^{20}\) used the original distance and Jaccard distance to complete re-ranking. Jiang et al.\(^{21}\) used spatial–temporal relationship among vehicles to re-rank the initial ranking list. However, in the vehicle re-ID task, vehicle images with orientations similar to that of the query image often occupy a considerable portion of the top-\(k\) returns in the ranking list. Therefore, use of the top-\(k\) results to enhance the features directly may not contribute substantially to optimizing the retrieval rankings of vehicle images with orientations dissimilar to that of the query image. Hence, we propose an orientation-guide query expansion method to optimize the rankings of images whose orientations are dissimilar to that of the query image by adding vehicle orientation information to the re-ranking process, thereby improving the retrieval performance.

The proposed orientation-guide query expansion approach consists of three steps. First, we extract image features using a feature extractor and select appropriate similarity measures to calculate the similarity matrix. Here, the vehicle orientation information can be obtained based on methods such as manual annotation and classifier prediction. Then, based on the similarity matrix and vehicle orientation information, we calculate the nearest neighbor image with the highest similarity and dissimilar orientation to the query image. Finally, the new features are obtained by fusing the feature of the query and nearest neighbor image obtained in the previous step, and image retrieval is performed again for the fused data. In the feature fusion process, we weight the original and nearest neighbor features based on their similarities. The contributions of this study are summarized as follows:

- An orientation-guide query expansion method is proposed to optimize the rankings of images that are dissimilar to the query image by including vehicle orientation information in the query expansion process.
- A feature fusion method based on similarity is used to reduce the influences of unexpected samples on the features, and the weights of the original features are increased to reduce the impacts on the retrieval of images whose orientations are similar to that of the query image.
- Extensive experiments are conducted using two public datasets to show that the proposed method can achieve a relative performance improvement.

**Related work**

**Orientation-aware vehicle Re-ID**

The vehicle re-ID task has garnered increasing attention in the field of computer vision. VehicleID\(^{22}\) and VeRi-776\(^{23}\) are two widely used datasets for vehicle re-ID tasks. In these datasets, any two images of the same vehicle may have different orientations, thereby affecting vehicle recognition. Therefore, Wang et al.\(^{24}\) used an orientation invariant feature embedding scheme to extract the features of the vehicle from different orientations based on 20 key-point locations. Chu et al.\(^{25}\) proposed a viewpoint-aware network that learns two kinds of metrics in two feature spaces for similar and different orientations. Meng et al.\(^{26}\) proposed a parsing-based view-aware embedding network to achieve view-aware feature alignment and enhancement for vehicle re-ID. Sun et al.\(^{27}\) used orientation information to learn two different metrics according to whether there is common field of view of two vehicle images to deal with large intra-class difference.
Re-ranking for re-ID

Re-ranking is a method used to optimize the ranking of the images, which can effectively improve retrieval performance. Re-ID can be regarded as a retrieval problem; some methods of image retrieval re-ranking can also be used in the re-ID task. In particular, Jegou et al.\(^\text{17}\) proposed the CDM that considers the neighborhood of a point. Arandjelović and Zisserman\(^\text{28}\) proposed the discriminative query expansion (DQE) method where a richer model for the query is learnt discriminatively in a form suited to immediate retrieval via use of the inverted index. Bai and Bai\(^\text{19}\) proposed a SCA scheme to encode the neighbor set into a vector and indicate sample similarities based on the generalized Jaccard distance.

Recently, some re-ranking methods have been proposed for the person re-ID tasks. For instance, Zhong et al.\(^\text{20}\) proposed a re-ranking method with \(k\)-reciprocal encoding, which encodes the \(k\)-reciprocal features into a single vector to obtain the \(k\)-reciprocal features of an image. Li et al.\(^\text{29}\) developed a re-ranking model by analyzing the relative information and direct information of the nearest neighbors of each pair of images. Garcia et al.\(^\text{30}\) refined a given initial ranking by removing the visual ambiguities common to the first ranks by analyzing their content and contextual information. Chen et al.\(^\text{31}\) incorporated graph models into feature subsets resorting to the initial ranking in the graph convolution network by integration of the attention mechanism. In addition, in the vehicle re-ID task, some studies\(^\text{21,24}\) have used spatial–temporal information to re-rank the initial list obtained with visual features. Shi et al.\(^\text{32}\) perform hash learning by calculating the semantic similarity among seen classes. However, owing to the extreme orientation variations in vehicle images, objects that need to be primarily optimized are the vehicle images whose orientations are dissimilar to that of the query image. Inspired by this idea, in this study, we designed a simple but effective re-ranking method based on query expansion and orientation information.

Method

This proposed orientation-guide query expansion approach consists of three steps. As illustrated in Figure 2, given a query image of a vehicle, first, we extract image features from the Reid Model, and select the appropriate similarity measure to calculate the similarity matrix, and obtain an initial ranking list. Wherein, the vehicle orientation information can be obtained based on some methods, such as manual annotation and classifier prediction. Then, according to the similarity matrix and vehicle orientation information, we calculate the nearest neighbor image with the highest similarity and dissimilar orientations to the query image. Finally, the new feature is obtained by fusing the feature of the query image and the nearest neighbor image obtained in the previous step, which is used to retrieve again for the final result. In the process of feature fusion, we weight the original feature and the nearest neighbor feature based on similarity.

Problem statement

We first extract the embedded features of the vehicle images in the database through the feature extractor, and then adopt the feature representation to get the initial similarity matrix and the ranking list according to the similarity measure. We define the query set
Table 1. The rule to determine whether two vehicle images have similar orientation according to the orientation label in the VeRi-776 data set. Where $O1 \sim O8$ represent “front,” “rear,” “left,” “left-front,” “left-rear,” “right,” “right-front,” and “right-rear,” respectively. And $S$ and $D$ represent two images in similar orientation and dissimilar orientation, respectively.

|    | 01 | 02 | 03 | 04 | 05 | 06 | 07 | 08 |
|----|----|----|----|----|----|----|----|----|
| 01 | S  | D  | D  | S  | D  | D  | S  | D  |
| 02 | D  | S  | D  | D  | S  | D  | D  | S  |
| 03 | D  | D  | S  | S  | S  | S  | S  | S  |
| 04 | S  | D  | S  | S  | D  | S  | S  | D  |
| 05 | D  | S  | S  | D  | S  | D  | S  | S  |
| 06 | D  | D  | S  | S  | S  | S  | S  | S  |
| 07 | S  | D  | S  | S  | D  | S  | S  | D  |
| 08 | D  | S  | S  | D  | S  | D  | S  | D  |

$Q = \{v^i_1 | i = 1, 2, \ldots, M\}$ with $M$ images and gallery set $G = \{v^i_2 | i = 1, 2, \ldots, N\}$ with $N$ images. $O(v_1, v_2)$ is used as the function to judge whether the two vehicles are in the similar orientation, where 0 and 1 represent that two vehicles are in the dissimilar orientation and the similar orientation, respectively.

For a query vehicle $v^q$, the initial ranking list $List(v^q, G) = \{v_1^q, v_2^q, \ldots, v_M^q\}$ can be obtained according to similarity matrix, wherein the samples are arranged in order of similarity. Our method aims to optimize the ranking in $List(v^q, G)$ of vehicle images whose orientation is dissimilar to the query image (i.e., a vehicle $v^q$ which satisfies $O(v^q, v^q) = 0$) to improve the performance of vehicle re-ID.

**Orientation-guide nearest neighbor search**

Our approach is based on vehicle orientation information. These orientation information are already existed in some datasets, such as the dataset VeRi-776. In VeRi-776, all vehicle images (including the training set and the test set) are manually labeled as eight types of orientations (“front,” “rear,” “left,” “left-front,” “left-rear,” “right,” “right-front,” and “right-rear”) by Wang et al.\textsuperscript{24} We directly use the label on the test set and add orientation information to the re-ranking operation. We develop a rule to determine whether two vehicle images have the similar orientation as shown in Table 1. The formulation of this rule is mainly based on whether there are common fields of view in two orientations, such as two pictures with the orientation of front and rear, we think their orientation is dissimilar because there are very few common fields of view. Conversely, for orientations of left and left-front, we consider their orientations are similar because there exists more common fields of view.

For some datasets that are lack of orientation information for all samples, such as the dataset VehicleID, which has only labeled the orientation information of the training set, we use ResNet50 as the backbone to train an orientation classifier to predict the orientation label of the test set. There are two orientations in VehicleID dataset, containing front and rear. When the orientation labels of two pictures are the same, the orientation is similar, otherwise, the orientation is dissimilar.

We get the similarity matrix of all images in the test set by cosine similarity. The similarity $s(v_1, v_2)$ between images $v_1$ and $v_2$ is described as

$$s(v_1, v_2) = \frac{f(v_1) \cdot f(v_2)}{\|f(v_1)\| \times \|f(v_2)\|},$$

where $f(\bullet)$ is the mapping function of the feature extractor, which maps image to feature vectors. According to the similarity matrix, we can get the initial ranking list of all the images. For a vehicle image $v$, its initial ranking list is $List(v) = (v_1^v, v_2^v, \ldots, v_M^v)$. and we look for the image $v'$ in $List(v)$ which has the highest similarity with $v$ and whose orientation is dissimilar to $v$, which is described as

$$v' = \arg \max_{v'} s(v, v')$$

$$s.t. O(v, v') = 0, v' \in List(v)$$

We will use features of $v$ and $v'$ to perform feature fusion operation.

**Feature fusion**

In the previous section, we described the orientation-guide nearest neighbor, $v'$, for a vehicle image, $v$, and combined the features of $v'$ and $v$ to obtain new features. In the AQE, the original and neighboring features are averaged to obtain the fused features; however, this method is not suitable for the orientation-guide query expansion because the rankings of vehicle images with similar orientations to the query image will be impacted if the original and neighboring features are fused with the same importance, although the rankings of vehicle images with dissimilar orientations to the query image would be set by cosine similarity.
may be optimized, thus resulting in overall performance degradation. Therefore, in the feature fusion process, we assign larger weights to the original features to reduce the impact on similar orientation image retrieval. Meanwhile, \( v_0 \) could be the sample retrieved by mistake. Adding the wrong sample into the fusion data will negatively impact the original features. When the retrieval is wrong, we consider that the similarity between the error sample and query image will be relatively low, so we weight the nearest neighbor features based on their similarity. Thus, the final formulation can be described as follows

\[
f_{\text{new}}(v) = \frac{f(v) + \lambda s(v, v')f(v')}{1 + \lambda s(v, v')} \quad (3)
\]

\[
f_{\text{new'}}(v) = \frac{f(v) + f(v')}{2} \quad (4)
\]

where \( \lambda \) is a weighted parameter used to reduce the proportion of nearest neighbor features; \( f_{\text{new}}(\bullet) \) and \( f_{\text{new'}}(\bullet) \) are the new features obtained after fusion. When there is weight, the original feature and the weighted neighbor feature are fused with new feature, that is, in equation (3). When there is no weight, the original feature and the initial nearest neighbor feature are directly fused with new feature, that is, in equation (4), the parameter \( \lambda \) and the cosine similarity \( s(v, v') \) are not included. We conducted the aforementioned operations on all images in the test set and obtained the new features of all the images. These new features were then adopted to calculate the new similarity matrices and the final ranking list of the images in \( Q \).

**Experiment**

In this section, we first introduce the experimental settings including two public benchmark sets used for vehicle re-ID and the evaluation metrics. Then, we compare our proposed method with several existing re-ranking methods based on the multiple feature extractor.

**Experimental settings**

In the experiments, we adopted two public datasets VeRi-776\(^2\) and VehicleID\(^2\) to vehicle re-ID, which are composed of vehicle images from the real-world surveillance video.

VeRi-776 contains over 50,000 images of 776 vehicles captured using 20 surveillance cameras in unconstrained traffic scenes. Among them, about 37,778 images of 576 vehicles are used for training, while the remaining 11,579 images of 200 vehicles are used for testing. VehicleID contains 221,763 images of 26,267 vehicles, which were captured during daytime by multiple real-world surveillance cameras distributed across a small city in China. VehicleID provides three test subsets with different sizes, namely, small, medium, and large, including 800, 1600, and 2400 vehicles, respectively.

We evaluated the proposed method based on two evaluation metrics, the mean average precision (mAP) and cumulative match characteristic (CMC) score. The former one is a widely used evaluation metric in retrieval tasks, which is considered as the mean value of the average precision (AP) for all the query images in this study. The latter one shows the probability of a query identity appearing in candidate lists with different sizes. In VeRi-776, we used mAP and CMC scores (rank-1 denoted as CMC@1 and rank-5 denoted as CMC@5)

### Table 2. Results (%) on VeRi-776 dataset, and the best results are shown in bold.

| Method          | mAP  | CMC@1 | CMC@5 |
|-----------------|------|-------|-------|
| Baseline        | 79.72| 95.53 | 98.45 |
| Baseline + \( k \)-reciprocal | 82.39| 97.02 | 97.74 |
| Baseline + AQE (\( k = 1 \)) | 81.37| 96.01 | 98.21 |
| Baseline + Ours | 85.24| 94.70 | 98.03 |
| Baseline + AQE (\( k = 9 \)) | 85.85| 96.78 | 97.56 |
| Baseline + Ours + AQE (\( k = 9 \)) | 89.76| 96.19 | 97.62 |

* mAP: mean average precision; CMC: cumulative match characteristic; AQE: average query expansion.

### Table 3. Results (%) on VehicleID dataset, and the best results are shown in bold.

| Method          | Small | Medium | Large |
|-----------------|-------|--------|-------|
|                 | CMC@1 | CMC@5  | CMC@1 | CMC@5  | CMC@1 | CMC@5  |
| Baseline        | 83.93 | 96.03  | 80.60 | 93.68  | 78.09 | 92.20  |
| Baseline + \( k \)-reciprocal | 82.31| 95.89 | 80.08 | 93.13  | 77.86 | 91.40  |
| Baseline + AQE (\( k = 1 \)) | 85.33| 96.56 | 82.06 | 94.46  | 79.01 | 93.15  |
| Baseline + Ours | 88.51| 97.52 | 85.89 | 95.93  | 83.60 | 94.87  |
| Baseline + AQE (\( k = 5 \)) | 85.00| 97.21 | 82.26 | 94.87  | 79.35 | 93.32  |
| Baseline + Ours + AQE (\( k = 5 \)) | **89.97**| **98.23**| **86.31**| **97.02**| **85.18**| **95.85**|

* CMC: cumulative match characteristic; AQE: average query expansion.
Specifically, we first set \( k = 1 \) because our method only fuses one nearest neighbor feature; then, we found the best \( k \) (9 for VeRi-776 and 5 for VehicleID) value for AQE. The results are shown in Tables 2 and 3.

As seen in Table 2, the mAP of our method outperforms the \( k \)-reciprocal and AQE (\( k = 1 \)) schemes. However, compared with the baseline, CMC@1 and CMC@5 have poor performances because the CMC scores focuses only on the first correctly retrieved sample, but our method has a negative impact on images whose orientations are similar to that of the query image. Generally, the first correctly retrieved sample is the image whose orientation is most similar to the query image. The mAP of AQE (\( k = 9 \)) is slightly larger than that of our method. However, using our method and AQE (\( k = 9 \)) together, we can obtain optimal performance. This shows that the samples targeted by these two methods are generally different. Besides, using our method first and then the AQE will further improve the performance index. Similarly, as shown in Table 3, our method outperforms other methods, and using our method with the AQE for the best \( k \) parameters can further improve performance.

### Table 4. Results (%) of other models on VehicleID dataset, and the best results are shown in bold.

| Method                  | Small CMC@1 | Small CMC@5 | Medium CMC@1 | Medium CMC@5 | Large CMC@1 | Large CMC@5 |
|-------------------------|-------------|-------------|--------------|--------------|-------------|-------------|
| BDBNet                  | 81.38       | 95.87       | 76.71        | 94.33        | 73.17       | 91.87       |
| BDBNet + k-reciprocal   | 82.47       | 96.26       | 76.50        | 90.29        | 73.68       | 88.00       |
| BDBNet + AQE (\( k = 1 \)) | 83.12       | 96.65       | 77.29        | 95.23        | 74.18       | 92.88       |
| BDBNet + Ours           | 83.28       | 96.65       | 77.43        | 91.59        | 75.13       | 89.34       |
| BDBNet + Ours + AQE (\( k = 5 \)) | 85.16       | 97.31       | 79.72        | 93.44        | 77.30       | 92.00       |
| PCB                     | 81.42       | 96.17       | 77.43        | 91.59        | 75.13       | 89.34       |
| PCB + k-reciprocal      | 82.47       | 96.26       | 76.50        | 90.29        | 73.68       | 88.00       |
| PCB + AQE (\( k = 1 \)) | 82.68       | 96.65       | 77.76        | 92.04        | 76.00       | 90.03       |
| PCB + Ours              | 85.81       | 97.31       | 82.70        | 94.97        | 80.66       | 93.45       |
| PCB + AQE (\( k = 5 \)) | 82.84       | 96.01       | 76.58        | 91.64        | 74.98       | 90.00       |
| PCB + Ours + AQE (\( k = 5 \)) | 87.91       | 98.26       | 82.58        | 94.79        | 80.81       | 93.79       |
| MGN                     | 85.16       | 97.93       | 79.72        | 93.99        | 76.52       | 91.68       |
| MGN + k-reciprocal      | 84.58       | 97.72       | 78.64        | 93.44        | 75.59       | 90.92       |
| MGN + AQE (\( k = 1 \)) | 85.42       | 98.45       | 80.08        | 94.40        | 77.30       | 92.00       |
| MGN + Ours              | 89.60       | 98.98       | 85.28        | 96.95        | 83.69       | 95.54       |
| MGN + AQE (\( k = 5 \)) | 84.88       | 97.40       | 79.09        | 94.30        | 76.88       | 91.51       |
| MGN + Ours + AQE (\( k = 5 \)) | 90.34       | 99.19       | 84.72        | 96.88        | 83.62       | 95.62       |

C: cumulative match characteristic; BDBNet: batch dropblock network; AQE: average query expansion; PCB: part-based convolutional baseline; MGN: multiple granularity network.

### Table 5. Results (%) of other models on VeRi-776 dataset, and the best results are shown in bold.

| Method                  | mAP | CMC@1       | CMC@5       |
|-------------------------|-----|-------------|-------------|
| PCB                     | 76.54 | 94.58       | 97.74       |
| PCB + k-reciprocal      | 79.17 | 95.17       | 96.36       |
| PCB + AQE (\( k = 1 \)) | 78.24 | 94.76       | 97.62       |
| PCB + Ours              | 83.83 | 92.43       | 96.96       |
| PCB + AQE (\( k = 9 \)) | 83.84 | 95.71       | 96.84       |
| PCB + Ours              | 87.85 | 94.52       | 96.48       |
| + AQE (\( k = 9 \))     |       |             |             |
| BDBNet                  | 79.44 | 96.01       | 98.45       |
| BDBNet + k-reciprocal   | 82.43 | 96.90       | 98.09       |
| BDBNet + AQE (\( k = 1 \)) | 81.43 | 96.48       | 98.33       |
| BDBNet + Ours           | 82.91 | 91.24       | 96.19       |
| BDBNet + AQE (\( k = 9 \)) | 86.12 | 96.54       | 98.15       |
| BDBNet + Ours           | 86.35 | 93.98       | 96.13       |
| + AQE (\( k = 9 \))     |       |             |             |
| MGN                     | 81.07 | 95.89       | 98.51       |
| MGN + k-reciprocal      | 83.08 | 96.90       | 97.56       |
| MGN + AQE (\( k = 1 \)) | 82.52 | 96.00       | 98.21       |
| MGN + Ours              | 87.23 | 93.86       | 97.74       |
| MGN + AQE (\( k = 9 \)) | 86.65 | 97.02       | 97.56       |
| MGN + Ours              | 90.46 | 95.65       | 97.08       |
| + AQE (\( k = 9 \))     |       |             |             |

mAP: mean average precision; CMC: cumulative match characteristic; PCB: part-based convolutional baseline; MGN: multiple granularity network; AQE: average query expansion; BDBNet: batch dropblock network.

**Experimental results and analyses**

We use a strong baseline (https://github.com/heshuting555/AICITY2020_DMT_VehicleReID)\(^{33-35}\) for the vehicle re-ID as the feature extractor. Based on the features, we applied our method and two frequently used re-ranking methods for re-ID, namely, AQE and \( k \)-reciprocal encoding. For our method, we set \( \lambda = 0.5 \). For the AQE, we used two different \( k \) parameters, where \( k \) is the number of fused nearest neighbor features. Specifically, we first set \( k = 1 \) because our method only fuses one nearest neighbor feature; then, we found the best \( k \) (9 for VeRi-776 and 5 for VehicleID) value for AQE. The results are shown in Tables 2 and 3.

As seen in Table 2, the mAP of our method outperforms the \( k \)-reciprocal and AQE (\( k = 1 \)) schemes. However, compared with the baseline, CMC@1 and CMC@5 have poor performances because the CMC scores focuses only on the first correctly retrieved sample, but our method has a negative impact on images whose orientations are similar to that of the query image. Generally, the first correctly retrieved sample is the image whose orientation is most similar to the query image. The mAP of AQE (\( k = 9 \)) is slightly larger than that of our method. However, using our method and AQE (\( k = 9 \)) together, we can obtain optimal performance. This shows that the samples targeted by these two methods are generally different. Besides,
To verify the universality of the vehicle re-ID method, in addition to a strong baseline, we added this method to several commonly used model architectures for re-ID, including the part-based convolutional baseline (PCB),36 multiple granularity network (MGN),37 and batch dropblock network (BDBNet).38 These results are shown in Tables 4 and 5.

As shown in Tables 4 and 5, we see that our method achieves reliable results for both databases and three models. It is noted that the CMC scores are a bit lower than some existing methods in some cases. However, the proposed method can well handle the vehicle images that are with different orientations with the query image, while the existing methods always focus on the correctly retrieved sample whose orientations are similar to that of the query image. Therefore, the whole performance (mAP score) of the proposed method is better than existing re-ranking methods, especially for the cases that are not similar with the query image in orientations.

A ss h o w ni nT a b l e s6 a n d7 , Baseline + Ours represents no weight and it is not equivalent to \( \lambda = 0 \), which means that there is no addition to the part based on similarity weight. Without weighting, performance will degrade a lot. Generally, large \( \lambda \) will have negative effect on the images whose orientation is similar to query

| Method            | Small        | Medium       | Large        |
|-------------------|--------------|--------------|--------------|
|                  | Rank-1 | Rank-5 | Rank-1 | Rank-5 | Rank-1 | Rank-5 |
| Baseline          | 83.93  | 96.03  | 80.60  | 93.68  | 78.09  | 92.20  |
| Baseline + Ours (\( \lambda = 1.5 \)) | 83.65  | 97.82  | 80.35  | 95.71  | 75.78  | 94.34  |
| Baseline + Ours (\( \lambda = 1 \))   | 86.84  | 97.87  | 84.03  | 96.05  | 81.77  | 94.91  |
| Baseline + Ours (\( \lambda = 0.8 \)) | 87.83  | 97.87  | 85.06  | 96.08  | 82.93  | 94.98  |
| Baseline + Ours (\( \lambda = 0.5 \)) | 88.51  | 97.52  | 85.89  | 95.93  | 83.60  | 94.87  |
| Baseline + Ours (\( \lambda = 0.3 \)) | 88.39  | 97.14  | 85.55  | 95.54  | 82.91  | 94.54  |
| Baseline + Ours (\( \lambda = 0.1 \)) | 86.49  | 96.49  | 82.91  | 94.44  | 80.45  | 93.43  |
| Baseline + Ours-nw | 74.11  | 97.08  | 70.43  | 93.81  | 65.14  | 91.94  |

Table 7. Results (%) of different parameter settings on VeRi-776 dataset, and the best results are shown in bold.

| Method            | mAP  | Rank-1 | Rank-5 |
|-------------------|------|--------|--------|
| Baseline          | 79.72| 95.53  | 98.45  |
| Baseline + Ours (\( \lambda = 1.5 \)) | 83.71| 90.76  | 94.87  |
| Baseline + Ours (\( \lambda = 1 \))   | 84.74| 91.78  | 95.83  |
| Baseline + Ours (\( \lambda = 0.8 \)) | 85.09| 92.43  | 96.66  |
| Baseline + Ours (\( \lambda = 0.5 \)) | 85.24| 94.70  | 98.03  |
| Baseline + Ours (\( \lambda = 0.3 \)) | 84.47| 95.65  | 98.21  |
| Baseline + Ours (\( \lambda = 0.1 \)) | 82.07| 95.95  | 98.51  |
| Baseline + Ours-nw | 83.07| 90.52  | 94.16  |

mAP: mean average precision.

Figure 3. The visualization of re-ranking results on VehicleID dataset.

Figure 4. The visualization of re-ranking results on VeRi-776 dataset.
image, and small $\lambda$ cannot well optimize the ranking of the images whose orientation is dissimilar to query image. In our experiment, the parameter $\lambda$ is set to 0.5.

The visualizations of the experimental results for the two data sets are shown in Figures 3 and 4. For each query image, the first line represents a part of the initial ranking list from the baseline model, and the second line depicts the results of re-ranking by our proposed method. It can be seen that our method optimizes the rankings of images whose orientations are dissimilar to that of the query image.

Conclusion

In this study, we aim to solve the problem of extreme orientation variation of vehicle image in the post-processing stage. By adding orientation information to the re-ranking operation, the proposed method can optimize the ranking of images whose orientation is dissimilar to query.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work is supported in part by the National Natural Science Foundation of China (62176141), Shandong Provincial Natural Science Foundation for Distinguished Young Scholars (ZR2021JQ26), Taishan Scholar Project of Shandong Province (tsqn202103088) and special funds for distinguished professors of Shandong Jianzhu University.

ORCID iD

Xue Zhang https://orcid.org/0000-0003-2429-101X

References

1. Klein LA, Kelley MR and Mills MK. Evaluation of overhead and in-ground vehicle detector technologies for traffic flow measurement. J Test Eval 1997; 25(2): 205–214.
2. Kuhne RD and Immes S. Freeway control systems for using section-related traffic variable detection. In: Proceedings of the Pacific Rim TransTech conference, Seattle, WA, 1993, https://cedb.asce.org/CEDBsearch/record.jsp?dockey=0083265
3. Balke KN, Ullman G, McCasland W, et al. Benefits of real-time travel information in Houston, Texas. Technical report, Texas Transportation Institute, Arlington, TX, 1995.
4. Woesler R. Fast extraction of traffic parameters and re-identification of vehicles from video data. In: Proceedings of the 2003 IEEE international conference on intelligent transportation systems, vol. 1, Shanghai, China, 12–15 October 2003, pp.774–778. New York: IEEE.
5. Shan Y, Sawhney HS and Kumar R. Vehicle identification between non-overlapping cameras without direct feature matching. In: Proceedings of the tenth IEEE international conference on computer vision (ICCV’05), vol. 1, Beijing, China, 17–21 October 2005, pp.378–385. New York: IEEE.
6. Guo Y, Rao C, Samarasekera S, et al. Matching vehicles under large pose transformations using approximate 3D models and piecewise MRF model. In: Proceedings of the 2008 IEEE conference on computer vision and pattern recognition, Anchorage, AK, 23–28 June 2008, pp.1–8. New York: IEEE.
7. Watcharapinchai N and Rujikietgumjorn S. Approximate license plate string matching for vehicle re-identification. In: Proceedings of the 2017 14th IEEE international conference on advanced video and signal based surveillance (AVSS), Leccce, 29 August–1 September 2017, pp.1–6. New York: IEEE.
8. Zapletal D and Herout A. Vehicle re-identification for automatic video traffic surveillance. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, Las Vegas, NV, 26 June–1 July 2016, pp.25–31. New York: IEEE.
9. Zheng Q, Liang C, Fang W, et al. Car re-identification from large scale images using semantic attributes. In: Proceedings of the 2015 IEEE 17th international workshop on multimedia signal processing (MMSP), Xiamen, China, 19–21 October 2015, pp.1–5. New York: IEEE.
10. He B, Li J, Zhao Y, et al. Part-regularized near-duplicate vehicle re-identification. In: Proceedings of the IEEE conference on computer vision and pattern recognition, Long Beach, CA, 15–20 June 2019, pp.3997–4005. New York: IEEE.
11. Khorramshahi P, Kumar A, Peri N, et al. A dual-path model with adaptive attention for vehicle re-identification. In: Proceedings of the IEEE international conference on computer vision, Seoul, Korea, 27 October–2 November 2019, pp.6132–6141. New York: IEEE.
12. Liu X, Zhang S, Huang Q, et al. Ram: a region-aware deep model for vehicle re-identification. In: Proceedings of the 2018 IEEE international conference on multimedia and expo (ICME), San Diego, CA, 23–27 July 2018, pp.1–6. New York: IEEE.
13. Guo H, Zhu K, Tang M, et al. Two-level attention network with multi-grain ranking loss for vehicle re-identification. IEEE T Image Process 2019; 28(9): 4328–4338.
14. Zhu J, Zeng H, Huang J, et al. Vehicle re-identification using quadruple directional deep learning features. IEEE T Intell Transp Syst, https://arxiv.org/abs/1811.05163
15. Yang J, Xing D, Hu Z, et al. A two-branch network with pyramid-based local and spatial attention global feature learning for vehicle re-identification. CAAI T Intell Technol 2021; 6(1): 46–54.
16. Chum O, Philbin J, Sivic J, et al. Total recall: automatic query expansion with a generative feature model for object retrieval. In: Proceedings of the 2007 IEEE 11th international conference on computer vision, Rio de Janeiro, Brazil, 14–21 October 2007, pp.1–8. New York: IEEE.
17. Jegou H, Schmid C, Harzallah H, et al. Accurate image search using the contextual dissimilarity measure. *IEEE T Pat Anal Mach Intell* 2008; 32(1): 2–11.
18. Leng Q, Hu R, Liang C, et al. Person re-identification with content and context re-ranking. *Multim Tools Appl* 2015; 74(17): 6989–7014.
19. Bai S and Bai X. Sparse contextual activation for efficient visual re-ranking. *IEEE T Image Process* 2016; 25(3): 1056–1069.
20. Zhong Z, Zheng L, Cao D, et al. Re-ranking person re-identification with k-reciprocal encoding. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, Honolulu, HI, 21–26 July 2017, pp.1318–1327. New York: IEEE.
21. Jiang N, Xu Y, Zhou Z, et al. Multi-attribute driven vehicle re-identification with spatial-temporal re-ranking. In: *Proceedings of the 2018 25th IEEE international conference on image processing (ICIP)*, Athens, 7–10 October 2018, pp.858–862. New York: IEEE.
22. Liu H, Tian Y, Yang Y, et al. Deep relative distance learning: tell the difference between similar vehicles. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, Las Vegas, NV, 27–30 June 2016, pp.2167–2175. New York: IEEE.
23. Liu X, Wu L, Ma H, et al. Large-scale vehicle re-identification in urban surveillance videos. In: *Proceedings of the IEEE international conference on multimedia and expo (ICME)*, Seattle, WA, 11–15 July 2016. New York: IEEE.
24. Wang Z, Tang L, Liu X, et al. Orientation invariant feature embedding and spatial temporal regularization for vehicle re-identification. In: *Proceedings of the IEEE international conference on computer vision*, Venice, 22–29 October 2017, pp.379–387. New York: IEEE.
25. Chu R, Sun Y, Li Y, et al. Vehicle re-identification with viewpoint-aware metric learning. In: *Proceedings of the 2019 IEEE/CVF international conference on computer vision*, Seoul, Korea, 27 October–2 November 2019, pp.8282–8291. New York: IEEE.
26. Meng D, Li L, Liu X, et al. Parsing-based view-aware embedding network for vehicle re-identification. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, Seattle, WA, 13–19 June 2020, pp.7103–7112. New York: IEEE.
27. Sun Z, Nie X, Xi X, et al. CFVMNet: a multi-branch network for vehicle re-identification based on common field of view. In: *Proceedings of the 28th ACM international conference on multimedia*, Seattle, WA, 12–16 October 2020, pp.3523–3531. New York: ACM.
28. Arandjelović R and Zisserman A. Three things everyone should know to improve object retrieval. In: *Proceedings of the 2012 IEEE conference on computer vision and pattern recognition*, Providence, RI, 16–21 June 2012, pp.2911–2918. New York: IEEE.
29. Li W, Wu Y, Mukunoki M, et al. Common-near-neighbor analysis for person re-identification. In: *Proceedings of the 2012 19th IEEE international conference on image processing*, Orlando, FL, 30 September–3 October 2012, pp.1621–1624. New York: IEEE.
30. Garcia J, Martinel N, Micheloni C, et al. Person re-identification ranking optimisation by discriminant context information analysis. In: *Proceedings of the IEEE international conference on computer vision*, Santiago, Chile, 7–13 December 2015, pp.1305–1313. New York: IEEE.
31. Chen X, Zheng L, Zhao C, et al. RRGCCAN: re-ranking via graph convolution channel attention network for person re-identification. *IEEE Access* 2020; 8: 131352–131360.
32. Shi Y, Nie X, Liu X, et al. Zero-shot Hashing via Asymmetric Ratio Similarity Matrix. *IEEE Transactions on Knowledge and Data Engineering* 2022.
33. He S, Luo H, Chen W, et al. Multi-domain learning and identity mining for vehicle re-identification. In: *Proceedings of the CVPR workshops*, 2020, https://arxiv.org/abs/2004.10547
34. Luo H, Jiang W, Gu Y, et al. A strong baseline and batch normalization neck for deep person re-identification. *IEEE T Multimedia*. Epub ahead if print 19 June 2019. DOI: 10.1109/TMM.2019.2958756.
35. Luo H, Gu Y, Liao X, et al. Bag of tricks and a strong baseline for deep person re-identification. In: *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR) workshops*, Long Beach, CA, 16–17 June 2019. New York: IEEE.
36. Sun Y, Zheng L, Yang Y, et al. Beyond part models: person retrieval with refined part pooling (and a strong convolutional baseline). In: Ferrari V, Hebert M, Sminchisescu C, et al. (eds) *Proceedings of the European Conference on Computer Vision (ECCV)*. Cham: Springer, 2018, pp.480–496.
37. Wang G, Yuan Y, Chen X, et al. Learning discriminative features with multiple granularities for person re-identification. In: *Proceedings of the 26th ACM international conference on multimedia*, Seoul, Korea, 22–26 October 2018, pp.274–282. New York: IEEE.
38. Dai Z, Chen M, Gu X, et al. Batch DropBlock network for person re-identification and beyond. In: *Proceedings of the IEEE international conference on computer vision (ICCV)*, Seoul, Korea, 27 October–2 November 2019. New York: IEEE.