Sample-Dependent Distance for 1 : N Identification via Discriminative Feature Selection

Naoki Kawamura and Susumu Kubota

Corporate Research and Development Center, Toshiba Corporation
Objective: Improving accuracy of identification

Collate a query with the registered N galleries

For the acceleration, feature vectors are often binarized in a high dimensional space in collating a query and galleries. We focus on the post-processing with the binary description.
Related Works

- Metric learning (e.g. Triplet loss*, Arcface loss**)

General metric learning modules are here

![Diagram showing metric learning process]

We focused on the post processing

Subject : Tackling to false-acceptance

Extracted future vectors potentially includes misleading factors even after deep metric learning in case where common features among different similar persons.

The misleading features should be ignored in identification.

The features appropriate for the identification would differ by samples.

The selected features should be determined for every sample.
We proposed a metric learning method as the feature selection, as the post-processing, for the binary description.

Approach: Sample-dependent feature selection

Feature extraction & Binary description

Similar feature patterns among the different identities are ignored.

The distance is bidirectionally calculated using the selected features depending on the samples.
Approach: Sample-dependent feature selection

- Select feature factors so that different but very similar identities can be well discriminated
- Each pair of a query and galleries are collated using the feature selection manner different by the samples

The distance is bidirectionally calculated using the selected features depending on the samples.
Approach: Sample-dependent feature selection

For each sample of a query and galleries, calculate feature selection manner

\[
\arg \min_{w(x_i)} \frac{1}{K} \sum_k \left[ (x_i \circ w(x_i))^T y_k^i \right]^2,
\]

Orthogonalization between a sample and the similar other identities

s.t. \[ \|w(x_i)\|_0 \leq R \]

Select \( R \) of \( D \) futures

\( x_i \in \{-1,1\}^D \): feature vector of the \( i \)-th sample

\( y_k(x_i) \in \{-1,1\}^D \): \( k \)-nearest gallery sample of the different identity of \( x_i \)

\( w(x_i) \in \{0,1\}^D \): feature selection mask of \( x_i \)
Sample-dependent feature selection is bi-directionally employed in collating a query $p$ and a gallery $g_i$.

- Hamming distance in usual cases:
  \[
  h(p, g_i) = \frac{D - p^T g_i}{2}
  \]
- Hamming distance under the feature selection of $p$:
  \[
  h_{w(p)}(p, g_i) = \frac{R - (w \circ p)^T g_i}{2}
  \]
- Proposed distance metric for the identification:
  \[
  h_w (p, g_i) = \frac{1}{2R} \left( h_{w(p)}(p, g_i) + h_{w(g_i)}(p, g_i) \right)
  \]
Experimental Results

• Comparison with the other methods of feature selection as the post-processing

• Comparable feature selections are applied to the outputs of various deep person-ReID models [***]

<table>
<thead>
<tr>
<th>Dataset</th>
<th># IDs (classes)</th>
<th># images</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Query</td>
<td>Gallery</td>
</tr>
<tr>
<td>Market1501</td>
<td>750</td>
<td>751</td>
</tr>
<tr>
<td>Dukemtmcreid</td>
<td>702</td>
<td>1110</td>
</tr>
</tbody>
</table>

Performances were stably improved using the proposed method

<table>
<thead>
<tr>
<th>Model</th>
<th>Data: train → test</th>
<th>Method</th>
<th>FRR@FAR [%] FAR=0.01%</th>
<th>FAR=0.1%</th>
<th>Rank1 [%]</th>
<th>mAP [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>OSNet</td>
<td>market→market</td>
<td>baseline</td>
<td>36.8</td>
<td>13.9</td>
<td>94.5</td>
<td>84.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proposed</td>
<td>29.7</td>
<td>10.1</td>
<td>95.0</td>
<td>86.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ReliefF</td>
<td>36.9</td>
<td>13.9</td>
<td>94.5</td>
<td>84.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fisher Score</td>
<td>37.1</td>
<td>13.9</td>
<td>94.6</td>
<td>84.5</td>
</tr>
<tr>
<td>OSNet</td>
<td>duke→duke</td>
<td>baseline</td>
<td>43.1</td>
<td>22.3</td>
<td>87.3</td>
<td>73.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proposed</td>
<td>36.0</td>
<td>19.2</td>
<td>88.3</td>
<td>75.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ReliefF</td>
<td>42.9</td>
<td>22.5</td>
<td>87.2</td>
<td>73.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fisher Score</td>
<td>43.3</td>
<td>22.3</td>
<td>87.4</td>
<td>73.4</td>
</tr>
<tr>
<td>OSNet</td>
<td>msmt→market</td>
<td>baseline</td>
<td>79.7</td>
<td>57.2</td>
<td>68.0</td>
<td>42.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proposed</td>
<td>77.1</td>
<td>53.6</td>
<td>69.4</td>
<td>42.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ReliefF</td>
<td>79.4</td>
<td>56.9</td>
<td>68.4</td>
<td>42.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fisher Score</td>
<td>80.2</td>
<td>58.0</td>
<td>68.3</td>
<td>42.3</td>
</tr>
<tr>
<td>OSNet</td>
<td>msmt→duke</td>
<td>baseline</td>
<td>62.3</td>
<td>40.4</td>
<td>67.2</td>
<td>49.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proposed</td>
<td>58.3</td>
<td>38.3</td>
<td>68.0</td>
<td>50.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ReliefF</td>
<td>62.0</td>
<td>40.2</td>
<td>67.5</td>
<td>49.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fisher Score</td>
<td>62.3</td>
<td>40.3</td>
<td>67.5</td>
<td>49.7</td>
</tr>
</tbody>
</table>

*** K. Zhou, et al., OSNet, ICCV2019
Experimental Results

Accuracy is also improved at very low FAR

The proposed method could discriminate quite a similar persons
Conclusion

• We focused on metric learning as the post-processing based on the feature selection

• Sample-dependent feature selection is employed so that each sample can be discriminated from the similar samples of different identities

• Our method stably improved the accuracy for a variety of 1:N-identification models