Person Re-identification using Co-occurrence Attributes of Physical and Adhered Human Characteristics

Masashi Nishiyama, Shota Nakano, Tappei Yotsumoto, Hiroki Yoshimura,
Yoshio Iwai and Kazunori Sugahara
Graduate School of Engineering, Tottori University, Japan
Email: nishiyama@eecs.tottori-u.ac.jp

***Abstract*—We propose a novel method for extracting features from images of people using co-occurrence attributes, which are then used for person re-identification. Existing methods extract features based on simple attributes such as gender, age, hair style, or clothing. Our method instead extracts more informative features using co-occurrence attributes, which are combinations of physical and adhered human characteristics (e.g., a man wearing a suit, 20-something woman, or long hair and wearing a skirt). Our co-occurrence attributes were designed using prior knowledge of methods used by public websites that search for people. Our method first trains co-occurrence attribute classifiers. Given an input image of a person, we generate a feature by vectorizing confidences estimated using the classifiers and compute a distance between input and reference vectors with a metric learning technique. Our experiments using a number of publicly available datasets show that our method substantially improved the matching performance of the person re-identification results, when compared with existing methods. We also demonstrated how to analyze the most important co-occurrence attributes.**

I. Introduction

Person re-identification is an active topic in pattern recognition research and has many potential applications such as watchlist monitoring and video surveillance. The problem is especially difficult because we must allow for variable viewpoints, illuminations, and poses. To overcome these difficulties, researchers have developed various approaches for extracting invariant features from images of people. These features have significant influences on the matching performance. In this paper, we focus on extracting invariant features that represent the people in images.

There are currently two main approaches for extracting features to be used in person re-identification applications. The first exploits low-level features [1]–[4] such as the distributions of gradients and colors. The second exploits high-level representations [5]–[8] such as gender, age, clothing, and gait, which are called human attributes in the field of soft-biometrics. However, it is exceptionally difficult to accurately infer high-level representations. Instead of inferring each attribute, researchers have recently started to use mid-level semantic attributes [9], [10] for person re-identification. These attributes directly represent elements of features extracted from images of people. However, existing methods for determining invariant features are not sufficient for person re-identification, when compared with characteristics that are used when people identify each other.

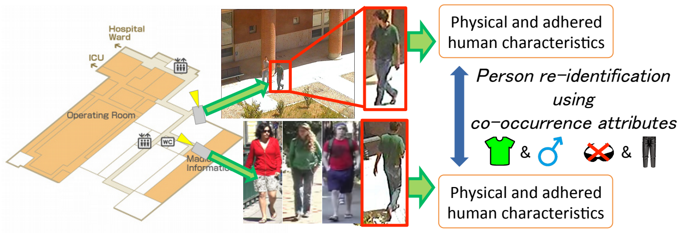

Fig. 1. Person re-identification using co-occurrence attributes. We propose using combinations of physical and human characteristics to extract features from images of people.

Our key idea is to exploit our prior knowledge of the characteristics used by people when looking for others. For instance, public websites of open criminal investigations [11], [12] or missing persons [13], [14] use a combination of physical and human characteristics such as gender and clothing. These combinations are considered to be informative clues when searching for people, and we expect that they would be useful invariant features for person re-identification.

In this paper, we attempt to determine what kind of combinations of physical and human characteristics are valuable for person re-identification, as illustrated in Figure 1. We call these combinations co-occurrence attributes. We propose a novel method for automatically extracting features based on the confidences of co-occurrence attributes, to enhance the performance of person re-identification methods. The extracted features were developed using our prior knowledge of the characteristics used in public websites that identify or find people. The rest of this paper is organized as follows. Section II contains a brief summary of related work, and Section III describes the design of the co-occurrence attributes. Section IV describes the proposed method for extracting features using co-occurrence attributes. Section V contains the results of our experiments and analyses, which demonstrate the methods effectiveness. Our concluding remarks are given in Section VI.

II. Related Work

Existing methods exploit low-level features such as histograms of oriented gradients [2], color histograms [1], or their combinations [3]. The learning-based approach in [4] selected effective low-level features from filter banks of gradients and colors. Unfortunately, low-level features are often affected

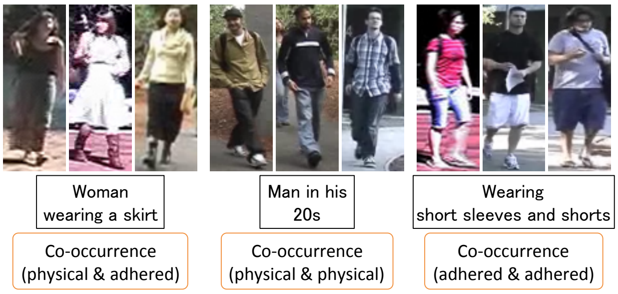

Fig. 2. Examples of co-occurrence attributes of physical and adhered human characteristics.

Fig. 3. Binary labels for representing co-occurrence attributes of physical (gender) and adhered human (tops) characteristics.

by the variability of viewpoints, illuminations, and poses. To extract invariant features, researchers have focused on human attributes [15]. We consider each characteristic (e.g., gender, hair style, or clothing) as a single attribute. Some methods [16], [17] improved the inference of single attributes using a joint learning technique. Khamis et al. [6] proposed jointly optimizing the inference of single attributes and the identification. Shi et al. [9] applied a domain shift between fashion and surveillance data to avoid complications due to labeling processes. Layne et al. [10] developed a method for weighting single attributes to increase the identification performance. However, existing methods have not sufficiently considered the combinations of physical and adhered human characteristics that are often used when people identify each other.

III. Design of co-occurrence attributes

A. Combinations of physical and human characteristics

We designed co-occurrence attributes that can be extracted from images of people. As described in [18], human attributes can be split into three intuitive types: physical, behavioral, and adhered human characteristics. Physical characteristics are person-specific traits and do not significantly change over time (e.g., gender, age, hairstyle, and beard). Behavioral characteristics are temporal changes such as gesture or gait. Adhered human characteristics depend on a persons appearance and are defined as things that are temporarily attached to a person (e.g., clothing or sunglasses).

We used combinations of physical and adhered human characteristics because we are motivated by certain public websites [11]–[14]. There are three ways to combine two characteristics, as illustrated in Figure 2: a combination of physical and adhered human characteristics (e.g., " woman wearing a skirt"), a combination of physical characteristics (e.g., "man in his 20s"); and a combination of adhered human characteristics (e.g., "wearing short sleeves and shorts").

B. Binary representation

Physical and adhered human characteristics use two types of class labels: binary labels such as gender (male or female) and sunglasses (presence or absence); and multi-class labels such as age (e.g., 20s, 30s, or over 40) and tops (e.g., long sleeves, short sleeves, or suit jacket). When simply combining two characteristics with L_1 and L_2 classes, the number of classes in the combination is $L_1 L_2$. Our method uses a binary label to represent a co-occurrence attribute by assuming that each characteristic class is independent. For instance, when combining gender ($L_1 = 2$) and tops ($L_2 = 3$), we obtain $L_1 L_2 = 6$ binary labels for the co-occurrence attributes. As illustrated in Figure 3, binary labels are represented as "man wearing long sleeves", "man wearing short sleeves", "man wearing a suit", "woman wearing long sleeves", "woman wearing short sleeves", and "woman wearing a suit" (and have true or false attributes). In biometrics, [18] showed that binary representations are intuitively easy for humans to understand, when compared with discrete and continuous representations. By exploiting the binary representation of co-occurrence attributes, our aim is to easily analyze which co-occurrence attributes are useful for person re-identification methods.

IV. Person re-identification using co-occurrence attributes

A. Overview of our method

We now briefly describe our method that uses co-occurrence attributes for person re-identification, which is illustrated in Figure 4. We train co-occurrence attribute classifiers to infer whether an image of a person contains the attributes. In this training process, we use images that have been labeled with co-occurrence attributes. These images are assigned positive or negative labels following the binary representation described in Section III-B. Note that this design complicates the labeling process, because there are many combinations of physical and adhered human characteristics. Thus, we automatically assigned co-occurrence attributes to the training samples using combinations of the labels for single attributes. If we have $L_1 L_2$ combinations of single attributes, there is 1 positive label of a co-occurrence attribute and $L_1 L_2 - 1$ negative labels. For instance, a positive label in Figure 3 is "man wearing long sleeves"; the remaining are negative.

Given an input image of a person, we compute the confidences of co-occurrence attributes using the trained classifiers. We generate a feature vector for person re-identification by vectorizing the confidences, and compute the distance between an input feature vector and a reference feature vector. The details of each step are described below.

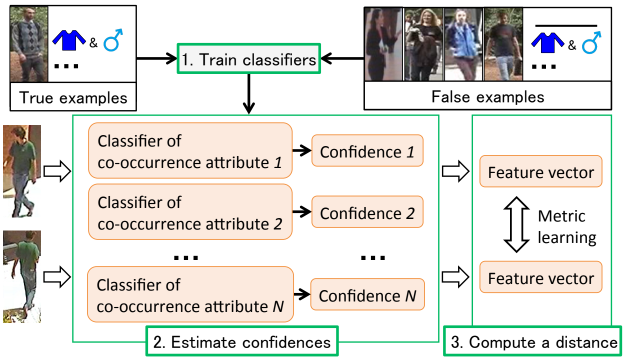

Fig. 4. Overview of our method that uses co-occurrence attributes for person re-identification.

B. Training classifiers of co-occurrence attributes

There may be much less positive attributes than negative. This is particularly the case when using co-occurrence attributes. If the number of positive and negative samples are significantly different, generic machine learning classification algorithms do not work well. To overcome the problems associated with imbalanced data, we can weight attributes according to inverse of the number of samples or align the number of samples. The alignment approach performed the best in our preliminary experiments, so this is the method we used in the remainder of this paper.

Next, we extracted feature vectors from the images to train the classifiers of co-occurrence attributes using the ensemble of localized features (ELF) method [4]. We used a linear support vector machine classifier SVM_i $(i \in 1, ..., N)$, where N is the number of co-occurrence attributes.

C. Extracting features for person re-identification

We describe how to compute the features used to match images of people. Given an image, we compute a signed distance h_i by applying a co-occurrence attribute classifier, that is,

$$h_i = \mathrm{SVM}_i(\boldsymbol{e}) \ , \tag{1}$$

where $\boldsymbol{e}$ represents an ELF vector extracted from an image of a person. An image is assigned a positive label when the signed distance is positive, and vice versa. We regard a signed distance as a confidence value, in the same manner as existing techniques. Note that if we directly use the signed distances as elements of a feature vector, the ranges of possible values are imbalanced because they depend on their respective element. Thus, we apply a simple scaling technique to align the ranges of values. We estimate the confidence value $x_i = 2h'_i - 1$ using

$$h'_i = \frac{h_i - \min h_i^t}{\max h_i^t - \min h_i^t} \ , \tag{2}$$

where $\max$ or $\min$ returns the maximum or minimum value h_i^t $(t \in 1, ..., T)$, and T is sum of the number of positive and negative samples. The feature vector for person re-identification is $\boldsymbol{x} = [x_1, x_2, ..., x_N]^{\mathrm{T}}$.

D. Computing the distance between feature vectors

We describe how to compute the distance between feature vectors extracted from images of people. If we use the all elements of a feature vector, we believe that some elements contribute to the identification and others do not. To increase the matching performance, some methods give larger weights to the important elements of a feature vector using a greedy algorithm technique [10] or a metric learning technique [19]. Our method used the large margin nearest neighbor (LMNN) method [20]. This technique generates a metric matrix $\mathbf{M}$ that reduces the distance between feature vectors belonging to the same person using the k-nearest neighbors, while lengthening the distance between vectors from different people. The distance, d, between features $\boldsymbol{x}_a, \boldsymbol{x}_b$ is

$$d^2 = (\boldsymbol{x}_a - \boldsymbol{x}_b)^{\mathrm{T}} \mathbf{M} (\boldsymbol{x}_a - \boldsymbol{x}_b) \ . \tag{3}$$

Note that d is smaller when the feature vector of individual a is more similar to the feature vector of individual b.

V. Experimental Analysis of Person Re-identification Using Co-occurrence Attributes

We evaluated the effectiveness of our method using some computational experiments. The person re-identification results are reported in Section V-A and our analysis of a metric matrix trained using LMNN is given in Section V-B. We also evaluated our method on a number of public datasets, as described in Section V-C.

A. Evaluation of basic performance

1) Experimental conditions: To evaluate the performance of our person re-identification method, we used the PETA Dataset [21], which consists of 10 publicly available datasets: 3DPeS [22], CAVIAR4REID [23], CUHK [24], GRID [25], i-LID [26], MIT [27], PRID [28], SARC3D [29], TownCentre [30], and VIPeR [31]. In the PETA dataset, all the images are labeled with single attributes. We selected 15 single attributes, as shown in Table I. P represents physical characteristics and A represents adhered human characteristics. In our experiments, we selected single attributes from [10] and public websites [11]–[14]. Note that we did not select color attributes, because colors significantly vary between surveillance cameras (as described in [18]).

We also designed 96 co-occurrence attributes, as shown in Table II. P&A represents combinations of physical and adhered human characteristics, P&P represents combinations of physical characteristics, and A&A represents combinations of adhered human characteristics. We removed 23 combinations of single attributes (e.g., "wearing a suit jacket and shorts") from the 119 combinations, because they did not commonly occur in practical cases. There were 59 P&A attributes, 22 P&P attributes, and 15 A&A attributes (Table II). The number of positive samples for each co-occurrence attribute was 48 or above.

We split the PETA dataset into test and training samples. We used images of people from VIPeR to evaluate the performance of the person re-identification method. We randomly selected 316 individuals from VIPeR for our test sample. The remaining 316 individuals were used to train a metric matrix for LMNN. We repeated the random selection 10 times to generate different test sample sets. We used two major indicators: cumulative match characteristic (CMC) curves dependent on the n-th rank matching rate, and the

TABLE I. SINGLE ATTRIBUTES FOR EVALUATING AN EXISTING METHOD. WE USED 2 CLASSIFIERS FOR BINARY CHARACTERISTICS AND 13 CLASSIFIERS FOR MULTI-CLASS CHARACTERISTICS.

	Binary label	Multi-class label
Physical characteristics (P)	Gender (Male/Female)	Hairstyle (Short/Long/Bald) Age (16–30/31–45/46–60/Over 60)
Adhered human characteristics (A)	Sunglasses (Presence/Absence)	Tops (Short Sleeves/Long Sleeves/Suit) Bottoms (Shorts/Skirt/Suit)

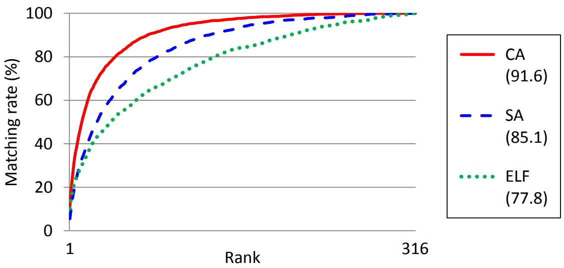

Fig. 5. CMC curve evaluated on the VIPeR dataset. The number in parentheses represents the nAUC for each method.

normalized area under the CMC curve (nAUC). We compared the following methods.

- **CA**: Our method for extracting features from images using the co-occurrence attributes described in Section IV.
- **SA**: An existing method [10] that extracts features using singles attributes. Note that we used the single attributes in Table I and LMNN instead of a greedy algorithm.
- **ELF**: An existing method [4] that extracts low-level features using filter banks of gradients and colors.

2) Experimental results: Figure 5 shows the performance of the person re-identification results based on the CMC curves and nAUC. The plot shows the average values for the 10 sets of randomly generated test samples. We can clearly see that our features based on the co-occurrence attributes are superior to features using singles attributes and the low-level features based on gradient and color. Overall, our method outperformed the others, and improved the matching rate for rank $n = 20$ by 20 points when applied to this difficult task.

To determine the most effective combinations of characteristics, we applied the method

- **P1**: without the combinations of physical and adhered human characteristics (P&A);
- **P2**: without the combinations of physical characteristics (P&P);
- **P3**: without the combinations of adhered human characteristics (A&A); and
- **P4**: using all the combinations in Table II.

TABLE II. CO-OCCURRENCE ATTRIBUTES FOR EVALUATING OUR METHOD.

Age 31–45 & Shorts (P&A)	Male & Long Hair (P&P)
Suit Jacket & Long Sleeves (A&A)	Age 46–60 & Short Sleeves (P&A)
Female &Age Above 60 (P&P)	Male & Age16–30 (P&P)
Short Hair & Suit Pants (P&A)	Bald & Short Sleeves (P&A)
Short Hair & Short Sleeves (P&A)	Female & Skirt (P&A)
Sunglasses & Short Hair (P&A)	Age31–45 & Shor tHair (P&P)
Age 46–60 & Suit Jacket (P&A)	Male & Long Sleeves (P&A)
Female & Suit Pants (P&A)	No Sunglasses & Shorts (A&A)
Sunglasses & Short Sleeves (A&A)	Age 46–60 & Bald (P&P)
Long Sleeves & Suit Pants (A&A)	No Sunglasses & Long Sleeves (A&A)
Age 31–45 & Long Hair (P&P)	Female & Age 31–45 (P&P)
Female & Suit Jacket (P&A)	Age 31–45 & Suit Pants (P&A)
Age 16–30 & Short Sleeves (P&A)	Female & Long Sleeves (P&A)
Short Hair & Skirt (P&A)	Male & Age Above 60 (P&P)
Male & Suit Jacket (P&A)	Age 16–30 & Suit Jacket (P&A)
Age 46–60 & Long Hair (P&P)	Male & Short Hair (P&P)
Female & Shorts (P&A)	Shorts & Short Sleeves (A&A)
Age 46–60 & Long Sleeves (P&A)	Sunglasses & Female (P&A)
Female & Age 46–60 (P&P)	Long Hair & Shorts (P&A)
Age 16–30 & Suit Pants (P&A)	Female & Short Hair (P&P)
No Sunglasses & Suit Pants (A&A)	Long Hair & Suit Pants (P&A)
Female & Age 16–30 (P&P)	Suit Jacket & Short Sleeves (A&A)
No Sunglasses & Suit Jacket (A&A)	Age 31–45 & Long Sleeves (P&A)
Skirt & Long Sleeves (A&A)	Sunglasses & Age 31–45 (P&A)
Age 16–30 & Shorts (P&A)	Age 16–30 & Short Hair (P&P)
Sunglasses & Long Sleeves (A&A)	Long Hair & Skirt (P&A)
Male & Short Sleeves (P&A)	Age 46– 60 & Short Hair (P&P)
Long Hair & Long Sleeves (P&A)	Age Above 60 & Short Sleeves (P&A)
Bald & Long Sleeves (P&A)	Age 31–45 & Short Sleeves (P&A)
Male & Suit Pants (P&A)	Female & Short Sleeves (P&A)
No Sunglasses & Age 46–60 (P&A)	Female & Long Hair (P&P)
Sunglasses & Male (P&A)	Suit Jacket & Suit Pants (A&A)
Age Above 60 & Bald (P&P)	No Sunglasses & Female (P&A)
No Sunglasses & Skirt (A&A)	Male & Age 31–45 (P&P)
Age Above 60 & Suit Jacket (P&A)	Age 46–60 & Suit Pants (P&A)
Short Hair & Suit Jacket (P&A)	Long Hair & Short Sleeves (P&A)
No Sunglasses & Age 31–45 (P&A)	Sunglasses & Age 16–30 (P&A)
Age Above 60 & Long Hair (P&P)	No Sunglasses & Age 16–30 (P&A)
Age 31-45 & Skirt (P&A)	Long Hair & Suit Jacket (P&A)
No Sunglasses & Short Hair (P&A)	Male & Age 46–60 (P&P)
Male & Bald (P&P)	No Sunglasses & Short Sleeves (A&A)
Male & Shorts (P&A)	Age 16–30 & Skirt (P&A)
Sunglasses & Long Hair (P&A)	No Sunglasses & Age Above 60 (P&A)
No Sunglasses & Bald (P&A)	Age 31–45 & Suit Jacket (P&A)
No Sunglasses & Long Hair (P&A)	Age Above 60 & Long Sleeves (P&A)
Age 16–30 & Long Hair (P&P)	No Sunglasses & Male (P&A)
Age 16–30 & Long Sleeves (P&A)	Short Hair & Long Sleeves (P&A)
Shorts & Long Sleeves (A&A)	Short Hair & Shorts (P&A)

Table III shows the matching rates and nAUCs for these four experiments. The combination of physical and adhered human characteristics increased the performance of the person re-identification method (the results for P1 were inferior to the others). The method using all the combinations (P4) produced the best results. These results demonstrate that the co-occurrence attributes of physical and adhered human characteristics developed using our method can successfully re-identify people in images.

TABLE III. COMPARISON OF MATCHING RATES AND NAUC TO EVALUATE THE EFFECTIVENESS OF THE DIFFERENT COMBINATIONS.

Plan	$n=1$	$n=5$	$n=10$	$n=20$	nAUC
P1	7.9	24.5	37.4	54.1	89.2
P2	9.9	29.5	42.1	58.6	90.9
P3	11.6	30.6	44.0	60.5	91.4
P4	11.3	32.0	45.6	63.9	91.6

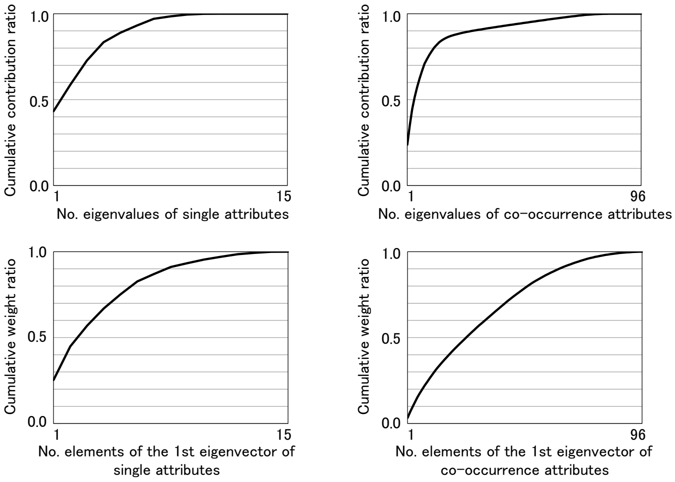

Fig. 6. Comparing the cumulative contribution ratio and the cumulative weight ratio curves between co-occurrence and single attributes.

B. Analysis of the metric matrix of LMNN

1) Analysis algorithm: To determine the most valuable co-occurrence attributes, we investigated a metric matrix $\mathbf{M}$ trained using LMNN. The matrix $\mathbf{M}$ is positive-semidefinite and is represented as

$$\mathbf{M} = \sum_{i=1}^{N} \lambda_i \boldsymbol{q}_i \boldsymbol{q}_i^{\mathrm{T}} , \quad (4)$$

where λ_i is the i-th eigenvalue and $\boldsymbol{q}_i$ is the i-th eigenvector. The eigenvectors corresponding to larger eigenvalues have the most important role when computing the distance between feature vectors (Equation (3)). We compute the cumulative contribution ratio C_m from λ_1 to λ_m using

$$C_m = \sum_{k=1}^{m} \frac{\lambda_k}{\sum_{i=1}^{N} \lambda_i} . \quad (5)$$

To determine the impact of the l-th element $q_{i,l}$ in eigenvector $\boldsymbol{q}_i$ when computing the distance between feature vectors, we compute the weight ratio $w_{i,l}$, i.e.,

$$w_{i,l} = \frac{|q_{i,l}|}{\sum_{j=1}^{N} |q_{i,j}|} , \quad (6)$$

and compute the cumulative weight ratio $W_{i,l}$ from $w_{i,1}$ to $w_{i,l}$, i.e.,

$$W_{i,l} = \sum_{k=1}^{l} w_{i,k} . \quad (7)$$

Note that $w_{i,k}$ $(k \in 1, ..., l)$ are in descending order.

TABLE IV. COMPARING THE IMPORTANCE OF ELEMENTS WHEN COMPUTING THE DISTANCE.

Co-occurrence attributes 1st eigenvector
Age 31–45 & Shorts (P&A)
Suit Jacket & Long Sleeves (A&A)
Short Hair & Suit Pants (P&A)
Short Hair & Short Sleeves (P&A)
Age 46–60 & Suit Jacket (P&A)
Female & Suit Pants (P&A)
Long Sleeves & Suit Pants (A&A)
Age 31–45 & Long Hair (P&P)

Co-occurrence attributes 2nd eigenvector
Short Hair & Shorts (P&A)
Female & Age 31–45 (P&P)
No Sunglasses & Age 31–45 (P&A)
Male & Age 46–60 (P&P)
Age 31–45 & Shorts (P&A)
Male & Suit Pants (P&A)
Age 16–30 & Long Hair (P&P)
Bald & Short Sleeves (P&A)

Co-occurrence attributes 3rd eigenvector
Age 16–30 & Suit Jacket (P&A)
Female & Skirt (P&A)
No Sunglasses & Suit Jacket (A&A)
Male & Age Above 60 (P&P)
Female & Age Above 60 (P&P)
Age 31–45 & Short Sleeves (P&A)

Single attribute 1st eigenvector
Short Hair (P)

2) Analysis results: We analyzed the matrix of co-occurrence attributes ($\mathbf{M}_{\mathrm{CA}}$) and the matrix of single attributes ($\mathbf{M}_{\mathrm{SA}}$). We investigated the elements of eigenvectors under the condition that the cumulative contribution ratios and cumulative weight ratios of $\mathbf{M}_{\mathrm{CA}}$ and $\mathbf{M}_{\mathrm{SA}}$ were nearly identical. Both matrices were trained using images of 632 individuals from the VIPeR dataset. We selected eigenvectors corresponding to the eigenvalues when $C_m = 0.4$: the 1st, 2nd and 3rd eigenvectors of $\mathbf{M}_{\mathrm{CA}}$ and the 1st eigenvector of $\mathbf{M}_{\mathrm{SA}}$. We selected elements of the eigenvectors for $W_{i,n} = 0.2$: eight elements of the 1st eigenvector of $\mathbf{M}_{\mathrm{CA}}$, eight elements of the 2nd eigenvector of $\mathbf{M}_{\mathrm{CA}}$, six elements of the 3rd eigenvector of $\mathbf{M}_{\mathrm{CA}}$, and one element of the 1st eigenvector of $\mathbf{M}_{\mathrm{SA}}$. Figure 6 shows the cumulative contribution ratio and the cumulative weight ratio curves of $\mathbf{M}_{\mathrm{CA}}$ and $\mathbf{M}_{\mathrm{SA}}$.

Table IV shows the informative elements of the eigenvectors of $\mathbf{M}_{\mathrm{CA}}$ and $\mathbf{M}_{\mathrm{SA}}$. There were 12 co-occurrence attributes containing age and seven co-occurrence attributes containing gender. Age and gender are indeed important for person re-identification, because these physical characteristics are also used in public websites [11]–[14]. There were six, three, and 13 co-occurrence P&P, A&A, and P&A attributes in Table IV, respectively. Three co-occurrence attributes contained short hair. These results show that an effective characteristic appears in both co-occurrence and single attributes. We believe that our method can effectively determine the characteristics used when people recognize each other.

C. Evaluation using public datasets

We evaluated the performance of our method using the 3DPeS, CAVIAR4REID, GRID, i-LID, PRID, SARC3D, and TownCentre datasets. We used test samples for person re-identification, training samples for LMNN, and training samples for the attribute classifiers with the experimental conditions described in Section V-A1. Table V shows the n-th matching rates and nAUC for the person re-identification results. We can see that our method using co-occurrence attributes is superior to the existing method [10] using single

TABLE V. COMPARING MATCHING RATE AND nAUC OF THE PERSON RE-IDENTIFICATION RESULTS USING A NUMBER OF PUBLIC DATASETS.

Dataset	Method	$n=1$	$n=5$	$n=10$	$n=20$	nAUC
3DPeS	**CA**	**20.0**	**39.8**	**51.0**	**65.0**	**78.6**
	SA	14.1	31.2	42.2	56.9	74.2
CAVIAR	**CA**	**24.7**	**55.8**	**72.3**	**90.1**	**81.7**
4REID	SA	15.6	40.7	59.2	80.2	72.9
GRID	**CA**	**18.9**	**46.4**	**65.1**	**80.3**	**89.9**
	SA	10.1	30.9	45.2	62.5	81.9
i-LID	**CA**	**20.9**	**46**	**60.7**	**77.6**	**81.2**
	SA	10.7	32.8	48.4	68.3	73.7
PRID	**CA**	**8.9**	**28.5**	**44.2**	**61.3**	**78.7**
	SA	4.1	19.5	31.4	50.6	71.3
SARC3D	**CA**	**60.9**	**94.5**	**98.9**	**99.5**	**96.5**
	SA	50.0	90.3	98.4	99.9	94.7
Town	**CA**	**32.3**	**49.9**	**57.8**	**67.6**	**82.5**
Centre	SA	18.4	35.1	44.2	55.1	74.6

attributes, on the all datasets. This significant improvement in the matching performance demonstrates the effectiveness of the co-occurrence attributes of combinations of physical and adhered human characteristics.

VI. CONCLUSION

We proposed a method for person re-identification that uses co-occurrence attributes. To extract invariant features from images of people, we introduced a combination of physical and adhered human characteristics. The contributions of this work can be summarized as follows.

- We designed co-occurrence attributes using physical and adhered human characteristics, which were based on prior knowledge of how people recognize each other.
- We analyzed a metric matrix trained using a metric learning technique to determine the most important co-occurrence attributes.

In the future, we will extend our method by considering combinations of three or more characteristics. We also intend to combine low-level features with our method.

REFERENCES

[1] C. Nakajima, M. Pontil, B. Heisele, and T. Poggio, "Full-body person recognition system," *Proceedings of Pattern recognition*, vol. 36, no. 9, pp. 1997–2006, 2003.

[2] X. Wang, G. Doretto, T. Sebastian, J. Rittscher, and P. Tu, "Shape and appearance context modeling," *Proceedings of International Conference on Computer Vision*, pp. 1–8, 2007.

[3] M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani, "Person re-identification by symmetry-driven accumulation of local features," *Proceedings of Computer Vision and Pattern Recognition*, pp. 2360–2367, 2010.

[4] D. Gray and H. Tao, "Viewpoint invariant pedestrian recognition with an ensemble of localized features," *Proceedings of European Conference on Computer Vision*, pp. 262–275, 2008.

[5] L. Bourdev, S. Maji, and J. Malik, "Describing people: A poselet-based approach to attribute classification," *Proceedings of International Conference on Computer Vision*, pp. 869–878, 2011.

[6] S. Khamis, C. Kuo, V. Singh, V. Shet, and L. Davis, "Joint learning for attribute-consistent person re-identification," *Proceedings of ECCV Workshop on Visual Surveillance and Re-Identification*, pp. 134–146, 2014.

[7] A. Li, L. Liu, K. Wang, S. Liu, and S. Yan, "Clothing attributes assisted person reidentification," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 134–146, 2015.

[8] J.W.Davis, "Visual categrization of children and adult walking styles," *Proceedings of ACM Multimedia*, pp. 295–300, 2001.

[9] Z. Shi, T. Hospedales, and T. Xiang, "Transferring a semantic representation for person re-identification and search," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4184–4193, 2015.

[10] R. Layne, T. Hospedales, and S. Gong, "Person re-identification," *chapter Attributes-Based Re-identification*, pp. 93–117, 2014, Springer.

[11] "Wanted," http://www.keishicho.metro.tokyo.jp/foreign/submenu.htm, Metropolitan Police Department.

[12] "Most wanted," http://www.scotland.police.uk/whats-happening/most-wanted/, Police Scotland.

[13] "Kidnappings & missing persons," https://www.fbi.gov/wanted/kidnap, Federal Bureau of Investigation.

[14] "Missing persons," http://www.scotland.police.uk/whats-happening/missingpersons/james-barr, Police Scotland.

[15] N. Kumar, A. Berg, P. N. Belhumeur, and S. Nayar, "Describable visual attributes for face verification and image search," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 10, pp. 1962–1977, 2011.

[16] L. Bourdev, S. Maji, and J. Malik, "Describing people: A poselet-based approach to attribute classification," *Proceedings of International Conference on Computer Vision*, pp. 1543–1550, 2011.

[17] J. Roth and X. Liu, "On the exploration of joint attribute learning for person re-identification," *Proceedings of 12th Asian Conference on Computer Vision*, pp. 673–688, 2014.

[18] A. Dantcheva, C. Velardo, A. Dfangelo, and J. Dugelay, "Bag of soft biometrics for person identification," *Multimedia Tools and Applications*, vol. 51, no. 2, pp. 739–777, 2011.

[19] S. Liao and S. Li, "Efficient psd constrained asymmetric metric learning for person re-identification," *Proceedings of International Conference on Computer Vision*, pp. 11–18, 2015.

[20] K. Weinberger, J. Blitzer, and L. Saul, "Distance metric learning for large margin nearest neighbor classification," *Journal of Machine Learning*, pp. 207–244, 2009.

[21] Y. Deng, P. Luo, C. Loy, and X. Tang, "Pedestrian attribute recognition at far distance," *Proceedings of ACM Multimedia*, pp. 3–7, 2014.

[22] D. Baltieri, R. Vezzani, and R. Cucchiara, "3dpes: 3d people dataset for surveillance and forensics," *Proceedings of joint ACM workshop on Human gesture and behavior understanding*, pp. 59–64, 2011.

[23] D. Cheng, M. Cristani, M. Stoppa, L. Bazzani, and V. Murino, "Custom pictorial structures for re-identification," *Proceedings of British Machine Vision Conference*, pp. 68.1–68.11, 2011.

[24] W. Ouyang and X. Wang, "A discriminative deep model for pedestrian detection with occlusion handling," *Proceedings of Computer Vision and Pattern Recognition*, pp. 3258–3265, 2012.

[25] C. Loy, T. Xiang, and S. Gong, "Time-delayed correlation analysis for multi-camera activity understanding," *International Journal of Computer Vision*, vol. 90, no. 1, pp. 106–129, 2010.

[26] T. Wang, S. Gong, X. Zhu, and S. Wang, "Person re-identification by video ranking," *Proceedings of European Conference on Computer Vision*, pp. 688–703, 2014.

[27] C. Papageorgiou and T. Poggio, "A trainable system for object detection," *International Journal of Computer Vision*, pp. 15–33, 2000.

[28] M. Hirzer, C. Beleznai, P. M. Roth, and H. Bischof, "Person re-identification by descriptive and discriminative classification," 2011, pp. 91–102.

[29] D. Baltieri, R. Vezzani, and R. Cucchiara, "Sarc3d: a new 3d body model for people tracking and re-identification," 2011, pp. 197–206.

[30] B. Benfold and I. Reid, "Stable multi-target tracking in real-time surveillance video," 2011, pp. 3457–3464.

[31] D. Gray, S. Brennan, and H. Tao, "Evaluating appearance models for recognition, reacquisition, and tracking," *Proceedings of Performance Evaluation of Tracking and Surveillance*, pp. 1–7, 2007.