The Journal of The Institute of Internet, Broadcasting and Communication (IIBC)
Vol. 17, No. 2, pp.15-23, Apr. 30, 2017. pISSN 2289-0238, eISSN 2289-0246

https://doi.org/10.7236/JIIBC.2017.17.2.15
JIIBC 2017-2-3

확장된 RNN을 활용한 사람재인식 시스템에 관한 연구

A Study on Person Re-Identification System using Enhanced RNN

최석규*, 허문걸**

Seok-Gyu Choi*, Wenjie Xu**

요 약 사람의 빈번한 자세 변화, 그리고 background clutter과 occlusion으로 인해 Person Re-identificatio는 컴퓨터 비전 분야에서 가장 어려운 부분이다. 비겹침 카메라의 이미지는 어떤 사람을 다른 사람과 구별하기 어렵게 한다. 더욱 나은 성능 일치를 달성하기 위해 대부분의 방법은 특징 선택과 거리 메트릭을 개별적으로 사용한다. 그렇게 차별화된 표현과 적절한 거리를 얻을 수 있고, 사람과 중요한 특징의 무시 사이의 유사성을 설명할 수 있다. 이러한 상황은 우리가 이 문제를 다루는 새로운 방법을 고려하도록 한다. 본 논문에서는 Person Re-identification를 위한 3단 계층 네트워크를 갖는 향상되고 반복적인 신경 회로망을 제안하였다. 특히 RNN(Revurrent Neural Network) 모델은 반복적인 EM(Expectation Maximum) 알고리즘과 3단 계층 네트워크를 포함하고, 차별적 특징과 지표 거리를 공동으로 학습한다. 반복적인 EM 알고리즘은 RNN 이전에 연속해 있는 CNN(Convoutional Neural Network)의 특징 추출 능력을 충분히 사용할 수 있다. 자율 학습을 통해 EM 프레임 워크는 패치의 레이블을 변경하고 더 큰 데이터 세트를 훈련할 수 있다. 네트워크를 더 잘 훈련시키기 위해 3단 계층 네트워크를 통해 CNN, RNN 및 풀링 계층이 공동으로 특징 추출을 할 수 있다. 실험 결과에 따르면 비전처리 분야에서 다른 연구자의 접근 방식과 비교할 때 이 방법은 경쟁력 있는 정확도를 얻을 수 있다. 이 방법에 대한 다른 요소의 영향은 향후 연구에서 분석되고 평가될 것이다.

Abstract The person Re-identification is the most challenging part of computer vision due to the significant changes in human pose and background clutter with occlusions. The picture from non-overlapping cameras enhance the difficulty to distinguish some person from the other. To reach a better performance match, most methods use feature selection and distance metrics separately to get discriminative representations and proper distance to describe the similarity between person and kind of ignoring some significant features. This situation has encouraged us to consider a novel method to deal with this problem. In this paper, we proposed an enhanced recurrent neural network with three-tier hierarchical network for person re-identification. Specifically, the proposed recurrent neural network (RNN) model contain an iterative expectation maximum (EM) algorithm and three-tier Hierarchical network to jointly learn both the discriminative features and metrics distance. The iterative EM algorithm can fully use of the feature extraction ability of convolutional neural network (CNN) which is in series before the RNN. By unsupervised learning, the EM framework can change the labels of the patches and train larger datasets. Through the three-tier hierarchical network, the convolutional neural network, recurrent network and pooling layer can jointly be a feature extractor to better train the network. The experimental result shows that comparing with other researchers' approaches in this field, this method also can get a competitive accuracy. The influence of different component of this method will be analyzed and evaluated in the future research.

Key Words : CNN, RNN, Unsupervised Learning, Person-Reidentification

*정회원, 전북대학교
**준회원, 전북대학교 컴퓨터공학부
접수일자: 2017년 3월 14일, 수정완료: 2017년 4월 4일
게재확정일자: 2017년 4월 7일

Received: 14 March, 2017 / Revised: 4 April, 2017 / Accepted: 7 April, 2017
*Corresponding Author: csg119@jbnu.ac.kr
Dept. Center for Advanced Image and Information Technology, School of Electronics & Information Engineering, ChonBuk National University, Korea

Ⅰ. Introduction

Human re-identification is a basic task in automated surveillance and has been a field of extensive research over the past years. Its aim is to distinguish the same person from the gallery according to a query picture. The challenge is that there are different cameras that monitor the same person, probing the same person from different source has some difficulty. Sometimes a camera may take hundreds of people in one day and some of them have a similar appearance. The same person under the different camera have significant challenges such as viewpoint variation, illumination, deformation, occlusion, background clutter, which make the person re-identification a challenge problem. Figure 1 shows some challenging examples in a most used dataset PRID_450S [15]. Each red box contains the matched person from the dataset PRID_450S.

그림 1. 이것은 데이터 집합 prid_450S에서 추출된 그림이며, 변형, 조명, 배경 클러스터에 따라 다르게 나온 동일 인물에 대한 인식 문제가 표시된 것이다. 빨간색 상자로 표시된 각 그룹은 각각 동일인에 대한 인식문제를 극복하는 대응 그림이다.

Fig. 1. this is the matched person from the dataset prid_450S, the challenge of viewpoint variation, illumination, background cluster are showed in the above picture, every group picture from the red box is the matched person which some kind of overcome these challenges.

There are traditional ways to predict the right patches form the gallery. Which use color model and weighted covariance estimation [17] to get the most similar part of the same person. Some traditional methods use algorithm to better get the salience to represent the patches. Our work is mainly manifested in two aspects. Given a query patch, the way to find correct person form a set of patches captured by the other camera must contain two important aspects. First, it must need a good image feature to represent the similar discriminative part of query and gallery patches. Good feature is necessary to the person re-identification due to it is the elemental unit of the picture. Second, it must need a fine-tuned distance metric to determine whether the gallery contains the query patch. This is critical because it determined the accuracy and similarity of the same person from the different person. Most methods use feature selection and distance metrics separately to get discriminative representations and proper distance to describe the similarity between people but some kind of ignoring some significant features or increasing the distance between the same people. State-of-art of the research has more focus on the first part and to get reliable and suitable feature to represent the person's visible appearance. Once get the good features from the dataset and next step is to get the suitable distance metrics. Some papers use a common distance such as L1 (Manhattan) or L2 (Euclidean) distance matrices to get the distance between patches.

This situation has encouraged us to consider a novel method to deal with this problem. Due to the good feature and distance metrics are the both critical part of the method, we use an iterative expectation maximum (EM) algorithm [16] which is contained in recurrent neural network (RNN). This novel EM based method automatically locates discriminative patches through utilizing the spatial relationships. In particular, we set that there is a variable patch form the dataset picture that state whether the feature is discriminative (the true good feature is descript by discriminative part of the picture). First of all, we set all the values into discriminative and then we train a RNN model that output the similarity of the same person according to the input data. We utilize the spatial smoothing to the final result similarity map and choose the patches that contain the most higher similarity and make it to be the real discriminative patches. We repeat this procedure by the true discriminative patches in the novel EM algorithm. Through the EM algorithm, the non-discriminative part is eliminated by the iterative process.

In our work we use enhanced recurrent neural network with three - tier hierarchical network for person re-identification. This ERNN-based architecture creates an inside circle to express the dynamic variation in the inputs, it not only can deal with the stationary patches but also the dynamic patches, so it is applicable on a broader field.

The main contributions of our proposed method is that we use the CNN architecture to better select the discriminative feature and use RNN to learning the distance metrics and join it together to be a group. To present a person's patch, the architecture can produce an output of 19950 dimension feature. Though the CNN and the last max pooling, the dimension can reduce. The proposed three-tier hierarchical network can learn the distance matric from a group of three pictures and this network make the same person's distance closer and far away the different person.

II. RELATED WORK

In the work of person re-identification, person re-ID based on still images and person re-ID based on video frames are two kinds of person re-identification.

Generally, there are two main categories on re-identification for still images: the method for extracting features from the input images usually called invariant feature representation, and the measure for comparing these features between images usually called distance metric learning. The first of methods uses a method based on invariant features that attempt to extract features that are distinguishable and invariant for changes in the environment and viewpoints.[21][34] Employs an ensemble of discriminant localized features and classifiers, which selected by contributing to improve invariance of viewpoint, combining spatial and color information. Kviatkovsky et al.[24] show that the color as a single cue under extremely variable imaging conditions has relatively good performance in identifying persons, and present a novel illumination invariant feature representation which is based on the log chromaticity (log) color space. The existing features include: local binary patterns[25, 11, 12, 20, 14] , salient color names [26], variations on color histograms[25, 11, 12, 20, 14, 29], Gabor features [14], and local patches [28].

Secondly, after extracting features, metric learning has been widely used for person re-identification to learn the metric emphasizing inter-personal distance and de-emphasizing intra-person distance. The basic idea of metric learning methods is to find the mapping from feature space to the new space in which eigenvectors from different image pairs are further than eigenvectors from same image pairs. The metric learning is applied to make a final decision as to if the person has been on the re-identified or not. [7] Presents Large margin nearest neighbor metric (LMNN) to improve the traditional KNN classification performance. Scale Invariant Local Ternary Pattern (SILTP) histograms [13] and a high dimensional color representation are proposed. While preserving local discrimination, it constructs the histogram of pixel features and then acquires its maximum values within horizontal band to handle the viewpoint variations.

Although it is often assumed in many ways for re-identification that a single image represents each person, in many real-world scenarios the use of video means that multiple images can be applied to enhance performance. Recently, performing person re-id in video sequences has been considered. Existing methods for video-based person re-id include Dynamic Time Warping (DTW)[31], conditional random field (CRF)[30], and top-push distance learning (TDL) Model [32][35].

The approaches of deep learning have also been exploited to video based person re-id to simultaneously solve the problem of feature representation and metric learning. In general, DNNs are applied to learn ranking functions based on image pairs[3] or triplets[33]. Such approaches typically depend on a deep network such as Siamese network[23], learning the direct mapping from original images to a feature space in which images from the identical person are approached while images from disparate person are widely separated.[27] Propose a “Siamese” convolutional network for metric learning. It includes three independent convolutional networks acting on three overlapping parts of the two input images. Each particular part of the network is composed of two convolutional layers with max pooling. Then the fully connected layer following, generates an output vector for each input image, and through a cosine function compare the two output vectors. The final similarity score are obtained by the cosine outputs for each of the three parts.[10] Propose two network : a recurrent neural network (RNN) and a Siamese network, one (RNN) is used to learn the interaction between multiple sequences in a video, the other is used to learn the distinguishing video-level features for person re-id. Long-Short Term Memory (LSTM) network[9] is proposed to assemble frame-wise person features in a recurrent way.

In this paper, the proposed enhanced recurrent neural network with three-tier hierarchical network for person use the CNN architecture to select the discriminative feature and use RNN to learning the distance metrics. The iterative EM algorithm can fully use the feature extraction ability of convolutional neural network (CNN) which is before the RNN. By unsupervised learning, the EM framework can change the labels of the patches and train larger datasets.

Ⅲ. PROPOSED METHOD

In this section, we descript the proposed methods in details. At the first, we will explain the whole neural network architecture about our person re-ID methods. Then we will specifically explain the proposed EM algorithm that joint with the RNN. Lastly, how can the proposed three-tier hierarchical network learn the distance matric from a group of three picture will be in detail.

1. The overall structure

As showed in figure 2, in order to train the network from the input images to get the most discriminative patches, we use triplet pictures as the input. Descript by the function of $\boldsymbol{T_m} = P(M,V) = \prod_{m=1}^{n}\prod_{k=1}^{n_m}(P(M_{m,k} \mid V_{m,k})P(V_{m,k}))$, there are three input images formed the m-th label, where the T_m^1 and T_m^2 are form the same person but T_m^3 are from the different person. As the other kind of CNN architecture[8], all the network share the same parameters W. we extracted the discriminative features from the raw images into a learned feature space. This is like $A_w(T_m)=(A_w(T_m^1),A_w(T_m^2),A_w(T_m^3))$. As a complete network, Each raw of the architecture from the figure1 is able to extract features from the training data. After the proposed network model is trained through the three-tier hierarchical network, we define that between $T_m^2 and T_m^3$ is more than the distance between $T_m^1 and T_m^2$, at the same time ,the D(T_m^1,T_m^2)< U(U is the defined number that limited the distance.). the three-tier hierarchical network can learning the distance from the training data make the same person distance near and the different person far away.

The Journal of The Institute of Internet, Broadcasting and Communication (IIBC)
Vol. 17, No. 2, pp.15-23, Apr. 30, 2017. pISSN 2289-0238, eISSN 2289-0246

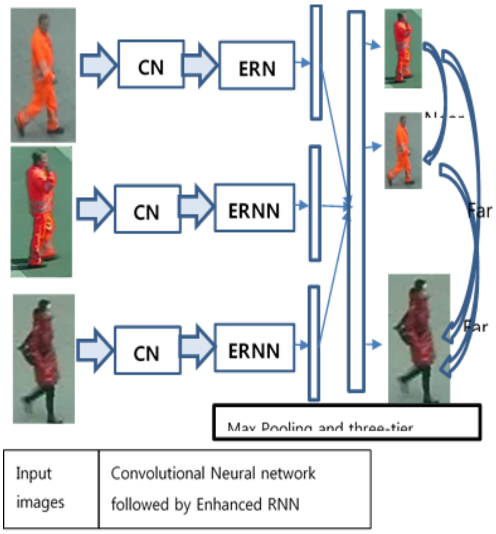

그림 2. 3-타이어 계층 네트워크를 갖춘 향상된 RNN
Fig. 2. An enhanced RNN with three-tire hierarchical network

2. The proposed recurrent neural network (RNN) model contain an iterative expectation maximum (EM) algorithm.

An overview of the EM are descript next. The dataset that we used is lower resolution images as patches that is extracted from the separate cameras. We have calculated the whole ground truth label that is contain the whole images not only some of the dataset. We decided the situation is just discriminative or non-discriminative as a binary tree.

We set that M = $\{M_1, M_2, M_n\}$ as the group that contains N bags. For every bag M_j = $\{M_{j,1}, M_{j,2}$, … $M_{j,N_j}\}$ have N_j situations. At the j-th bag, the j-th situation and relative label is in the $M_{i,j}=<M_{i,j}, y_j>$. Given that the bags are standalone and identically allocated, the M and the variables V can be calculated by the following model:

$$P(M,V)=\prod_{m=1}^{n}\prod_{k=1}^{n_m}(P(M_{m,k}\mid V_{m,k})P(V_{m,k})) \quad (1)$$

We maximize the patch similarity P (M) using the Expectation maximization.

First, at the beginning step, we should initial the E. we set $V_{i,j}$= 1 for all the i, j. From this set, we can consider all the situation is discriminative. In this step, we consider all the patch in a positive way to set the dataset.

At the maximization step, we make the patches similarity to become closer and make the parameter ε to get the higher possibility.

$$\begin{aligned}\varepsilon &\leftarrow \arg\max P(M\mid V;\varepsilon)\\ &=\arg\max\prod_{D}P(x_{m,k},y_m\mid\varepsilon)\\ &*\prod_{notD}P(x_{p,q},y_q\mid\varepsilon)\end{aligned} \quad (2)$$

Where the D represent the discriminative part of the patches. If we make a generative function for all the discriminative and latent instances, the optimization of the above function can be this:

$$\begin{aligned}&\arg\max\prod_{D}P(x_{m,k}y_m\mid x_{m,k};\varepsilon)\\ &=\arg\max\prod_{D}P(x_{m,k},y_m\mid\varepsilon)P(x_{m,k}\mid\varepsilon)\end{aligned} \quad (3)$$

Finally, we valued the hidden variables V. specially, $V_{m,k}$=1 is a certain threshold and it is initial at the first. As for the person re-ID, $P(M_{m,k}\mid V)$ is got by applying Gaussian smoothing at $P(x_{m,k}y_m\mid x_{m,k};\varepsilon)$, though this way, the smoothing step can utilizes the spatial Mutual position of $P(x_{m,k}y_m\mid x_{m,k};\varepsilon)$. At last, we repeat back to the Maximization step until it is become convergence.

Ⅳ. EXPERIMENT RESULTS

In this section, we evaluate our method on the most popular dataset, and perform a comprehensive evaluation of the proposed method. And then we compare different variations of the functions to compare the results. First, we describe these datasets and evaluation principle. Second, through the experiment, we predict some reliable results compare with some traditional methods and state-of -art method for some specific situation. Finally, we illustrate some disadvantages and advantages of this method and state the future work that we will do next.

1. Datasets PRID-450S.

To perform our experiments, we consider the most challenging datasets. The PRID 450S Dataset consists of 450 images pairs of pedestrians captured by two non-overlapping cameras. The main challenges are related to changes in viewpoint, pose as well as significant differences in background and illumination.

2. Evaluation principle

For the evaluation, we use the most useful cumulative match curve (CMC) metric for feature match. Towards the PRID_450S dataset, we choose half of it as the gallery and then choose the other half as the query for testing the network. We randomly choose the gallery and query to avoid some errors. For this dataset that we used, we use the camera A as the gallery and the picture from camera B are chosen to be the query. As for the multi-camera groups, we plan to choose two same person, one as the gallery and the other as the query. We make sure that each individual is included in the gallery dataset. When start to search the query picture from the gallery, the previous network should be executed first. The whole gallery picture is calculated through the network and get the discriminative feature from the result. Then, we compare the distance between the query picture and the whole gallery dataset using the L2 distance matric. Finally, we get the top-n nearest images. If the match picture is in the n nearest scale, the rank of the correct is rank n. In order to eliminate the deviation, we repeat this process ten times to get the mean number.

3. Experimental setting

Data augmentation : Similar to the references [1, 2], we use random partitions of the dataset and half of it as gallery and half it as the query. In order to eliminate deviations, we repeat this process ten times. In order to improve accuracy and increase training data, we use data augmentation before we apply the dataset. This is also conducive to prevent over-fitting problem. In practice, the original image is 100*220 pixels. But during the input process, we cut a region of 95*210 with a random shiver, finally we get the more reliable and abundant data to train. The result are get by the CMC curve by top n positions. If the match picture show up at the n-th position, then the result is rank n.

4. Feature Extraction

To represent the most discriminative part of the picture, we use the hand-craft label and ERNN followed by CNN at the same time. The dimension of the image is 19950, but after the CNN and ERNN, the dimension is reduced and get the relative feature.

5. Compared methods

To our best knowledge, our CNN+ERNN is the most novel method with better result compared with the other methods. We trust the better result comes from the combination of CNN and RNN with the three-tier architecture. Nevertheless, learning how to limit the matching features to the gallery from the non-overlapping camera is a very challenging task. But here we just do the experiment in the dataset of PRID_450S.

According to the results from table1, we completed the improved outputs while compared to efficient impostor-based metric learning [18], Large scale metric learning[22], Salient color names[19] and other approaches. For example, we reached much higher ranking rate comparing to KISSME method, this can be explained though the lower dimensional features and at the same time, same the calculating time.

표 1. PRID-450S 데이터 세트에 대한 실험 결과
Table 1. experimental result on PRID-450S dataset.

Methods	R=1	R=5	R=10	R=20	R=30
KISSME[15]	33.0	59.8	71.0	79.0	84.5
EIML[4]	35.0	58.5	68.0	77.0	83.0
KPCA[5]	42.9	67.7	76.6	84.7	89.5
KPLS[6]	40.5	63.1	71.6	81.3	87.2
KPLS Mode A	50.0	78.6	87.0	93.7	96.0
CNN+ERNN	52.8	81.3	90.0	94.0	96.2

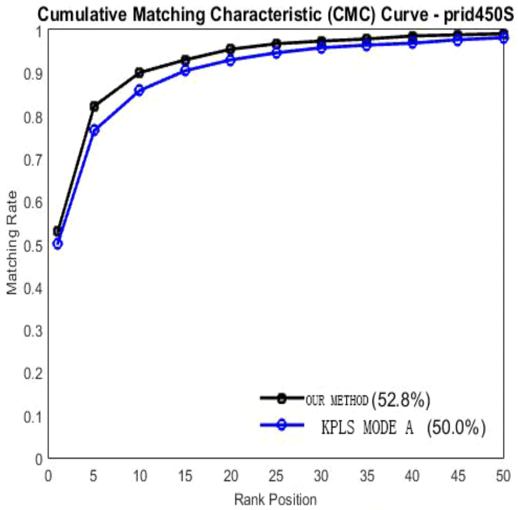

그림 3. KPLS 모드 A와 우리의 방법에 대한 최첨단 방법이 여기에 나와 있다. 이 방법은 다른 알고리즘과 비교 결과를 비교할 수 있다.
Fig. 3. The most state-of-art method of KPLS mode A and our method are showed in here. The method can get a comparative result compare the other algorithm.

The figure3 shows that CMC curve representation on the most state-of-art method of KPLS mode A and our method. This is applied on the person re- ID dataset PRID-450S. The contributions of our paper is that our proposed person re-ID architecture has three innovative part: 1) we use the CNN architecture to better select the discriminative feature and use RNN to learning the distance metrics and join it together to be a group. 2) the proposed EM algorithm that joint with the RNN can product patch-level outputs to see the discriminative part as a binary number, through the iteration, the most discriminative feature for person has been extracted. 3) the proposed three-tier hierarchical network can learn the distance matric from a group of three pictures and this network make the same person's distance closer and far away the different person.

V. CONCLUSIONS AND FUTURE WORKS

In this paper, we proposed an enhanced recurrent neural network with three-tier hierarchical network for person re-identification. Specifically, the proposed recurrent neural network (RNN) model contain an iterative expectation maximum (EM) algorithm and three-tier Hierarchical network to jointly learn both features and distance. The iterative EM algorithm can fully use of the feature extraction ability of convolutional neural network (CNN) which is in series before the RNN. Through the three-tier hierarchical network, the convolutional neural network, recurrent network and pooling layer can jointly be a feature extractor to better train the network. Comparing with the other result in this field, this method also can get a competitive accuracy. We will fine-tune the program to increase the accuracy and increase the compatibility to the other kind of dataset. Because RNN is an internal state of network that could express the dynamic temporal gesture, this makes it suitable for person re-ID in the video. This is a big step in the person tracking area also. We will explore this later and use this method run smoothly on the other kind of dataset.

References

[1] A. Globerson and S. T. Roweis. Metric learning by collapsing classes. In NIPS, pages 451 - 458, 2005.
[2] D. Gray and H. Tao. Viewpoint invariant pedestrian recognition with an ensemble of localized features. In ECCV, pages 262 - 275. 2008.
[3] D. Yi, Z. Lei, S. Liao, and S. Z. Li, "Deep metric learning for person re-identification," in 2014 22nd International Conference on Pattern Recognition (ICPR). IEEE, pp. 34 - 39, 2014.
[4] H. Martin, R. Pete, and B. Horst, "Person re-identification by efficient impostor-based metric learning," in IEEE AVSS, pp. 203 - 208, 2012.
[5] Raphael Prates and William Robson Schwartz. "

Kernel Hierarchical PCA for Person Re-Identification" in ICPR 2016.

[6] Raphael Prates, Marina Oliveira and William Robson Schwartz. " Kernel Partial Least Squares for Person Re-Identification" in AVSS 2016.

[7] K. Q. Weinberger and L. K. Saul, "Distance metric learning for large margin nearest neighbor classification," Journal of Machine Learning Research, vol. 10, no. Feb, pp. 207–244, 2009.

[8] De Cheng, Yihong Gong, Sanping Zhou, Jinjun Wang, Nanning Zheng. "Person Re-Identification by Multi-Channel Parts-Based CNN with Improved Triplet Loss Function" in CVPR 2016.

[9] Y. Yan, B. Ni, Z. Song, C. Ma, Y. Yan, and X. Yang, "Person identification via recurrent feature aggregation," in European Conference on Computer Vision. Springer, pp. 701–716, 2016.

[10] McLaughlin, N., Martinez del Rincon, J., & Miller, P. Recurrent Convolutional Network for Video-based Person Re-Identification" in CVPR 2016.

[11] S. Khamis, C. Kuo, V. Singh, V. Shet, and L. Davis. Joint learning for attribute-consistent person re-identification. In ECCV Workshop on Visual Surveillance and Identification, 2014.2.

[12] M. Koestinger, M. Hirzer, P. Wohlhart, P. Roth, and H. Bischof. Large scale metric learning from equivalence constraints. In CVPR, 2012. 2, 6.

[13] S. Liao, Y. Hu, X. Zhu, and S. Z. Li, "Person re-identification by local maximal occurrence representation and metric learning," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2197–2206, 2015.

[14] W. Li and X. Wang. Locally aligned feature transforms across views. In CVPR, 2013. 2

[15] M. Koestinger, M. Hirzer, P. Wohlhart, P. Roth, and H. Bischof, "Large scale metric learning from equivalence constraints," in IEEE CVPR, 2012.

[16] Le Hou, Dimitris Samaras, Tahsin M. Kurc, Yi Gao, James E. Davis, Joel H. Saltz; The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2424-2433, 2016.

[17] Yang Yang, Shengcai Liao, Zhen Lei, Dong Yi, and Stan Z. Li, "Color models and weighted covariance estimation for person re-identification," in IAPR ICPR, pp. 1874–1879, 2014.8.

[18]M. Hirzer, P.M. Roth, and H. Bischof, "Person re-identification by efficient impostor-based metric learning," in IEEE AVSS, p. 203–208, 2012.9.

[19] Yang Yang, Jimei Yang, Junjie Yan, Shengcai Liao, Dong Yi, and Stan Z Li, "Salient color names for person reidentification," in ECCV, pp. 536–551. 2014.

[20] N. Martinel, C. Micheloni, and G. Feresti. Saliency weighted features for person re-identification. In ECCV Workshop on Visual Surveillance and Re-identification, 2014. 2.

[21] D. Gray and H. Tao, "Viewpoint invariant pedestrian recognition with an ensemble of localized features," in European conference on computer vision. Springer, pp. 262–275, 2008.

[22] Martin Koestinger, Martin Hirzer, Paul Wohlhart, Peter M. Roth, and Horst Bischof, "Large scale metric learning from equivalence constraints," in IEEE CVPR, 2012.

[23] R. Hadsell, S. Chopra, and Y. LeCun, "Dimensionality reduction by learning an invariant mapping," in 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), vol. 2. IEEE, pp. 1735–1742, 2006.

[24] I. Kviatkovsky, A. Adam, and E. Rivlin, "Color invariants for person identification," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 7, pp. 1622–1634, 2013.

[25] F. Xiong, M. Gou, O. Camps, and M. Sznaier. Person reidentification using kernel-based metric learning methods. In ECCV, 2014. 2.

[26] Y. Yang, J. Yang, J. Yan, S. Liao, D. Yi, and S. Li. Salient color names for person re-identification. In

ECCV, 2014. 2.

[27] D. Yi, Z. Lei, and S. Z. Li. Deep metric learning for practical person re-identification. ICPR, 2014. 2, 5.

[28] Z. Zhang, Y. Chen, and V. Saligrama. A novel visual word co-occurrence model for person re-identification. In ECCV Workshop on Visual Surveillance and Re-identification, 2014. 2, 7.

[29] R. Zhao, W. Ouyang, and X. Wang. Person re-identification by salience matching. In ICCV, 2013. 2, 7.

[30] S. Karaman and A. D. Bagdanov, “Identity inference: generalizing person re-identification scenarios,” in European Conference on Computer Vision. Springer, pp. 443 - 452, 2012.

[31] D. Simonnet, M. Lewandowski, S. A. Velastin, J. Orwell, and E. Turkbeyler, “Re-identification of pedestrians in crowds using dynamic time warping,” in European Conference on Computer Vision. Springer, pp. 423 - 432, 2012.

[32] J. You, A. Wu, X. Li, and W.-S. Zheng, “Top - push video - based person re-identification,” arXiv preprint arXiv:1604.08683, 2016.

[33] S. Ding, L. Lin, G. Wang, and H. Chao, “Deep feature learning with relative distance comparison for person re-identification,” Pattern Recognition, vol. 48, no. 10, pp. 2993 - 3003, 2015.

[34] In-Kue Park, BoHyeok Ahn, Gyoo-Seok Choi, “An Edge Detection for Face Feature Extraction using λ- Fuzzy Measure”, The Journal of Internet Broadcasting and Communication(JIIBC), Vol.9, No.4, pp.69-74, August 2009.

[35] Kim, Tae-Hyun, “Measurement and Prediction of the Visibility Range by the Variations of the Character Sizes and Illuminance”, Journal of the Korea Academia - Industrial cooperation Society, Vol.16, No.12, pp.8222-8227, December 2015.
DOI : http://dx.doi.org/ 10.5762/KAIS. 2015.16.12. 8222.

저자 소개

최 석 규(정회원)

• 2000년 ~ 2006년 : 전북산업경영연구소 지역경제연구팀장(이사)
• 2006년 ~ : 전북대학교(연구교수, 교수)
<주관심분야 : 핀테크, 인공지능, ICT 융합>

허 문 걸(준회원)

• 2015년 : 중국 상하이대학교 학사
• 2016년 ~ : 전북대학교 컴퓨터 공학부 석사과정
• 2016년 ~ : Dept. Center for Advanced Image and Information Technology, School of Electronics & Information Engineering, ChonBuk National University
<주관심분야 : 인공지능, 로봇, 가상현실, 게임>

※ This paper was supported by research funds for newly appointed professors of Chonbuk National University in 2015