Method for Person Re-Identification with 2D-to-3D Image (Image-to-Video) Conversion

Kohei Arai
Science and Engineering Faculty, Saga University, Saga City, Japan

***Abstract*—A method for person re-identification using image-to-video conversion tools is proposed. The proposed method involves matching two images taken from different viewpoints: a reference image captured in advance and a current image captured in real-time to identify the person in concern in the current image, whose image is matched to the reference image. The 2D current image is converted into a 3D representation, from which synthetic images are generated at multiple viewpoints. By comparing the generated images and the reference image, person re-identification can be done. Experiments have demonstrated that the proposed method significantly improves identification accuracy. By accounting for changes in appearance due to different viewpoints and utilizing advanced image conversion, one of the main challenges in person re-identification is addressed. This approach offers a promising solution for applications requiring high accuracy in identifying individuals across varying perspectives.**

Keywords—Person re-identification; identification performance; 2D-to-3D image conversion method; TRIPO; CSM; KLING

I. Introduction

Person re-identification involves recognizing an individual across different images or video frames, typically within an anonymized dataset. This task is crucial in the fields of privacy protection and data security. However, existing methods often struggle with low accuracy because a person's appearance can vary significantly when viewed from different angles, and there may be many individuals with similar appearances in the dataset.

Challenges in person re-identification affecting recognition rates are as follows:

Person re-identification (Re-ID) is a crucial task in computer vision and surveillance, aiming to recognize and match individuals across different images or video frames, often captured in varying conditions and from non-overlapping camera views. Despite extensive research, achieving high recognition rates in person Re-ID remains challenging due to several factors:

1) Appearance variations:

a) Viewpoint differences: The same person can appear vastly different when viewed from various angles. Changes in posture, orientation, and perspective can significantly alter visual features.

b) Lighting conditions: Variations in lighting can affect the perceived color and texture of clothing, leading to inconsistent appearance across images.

c) Attire and accessories: Changes in clothing, addition or removal of items like jackets or bags, and seasonal attire shifts (e.g., winter coats versus summer wear) can make the same individual look different.

d) Personal belongings: The presence or absence of belongings such as umbrellas, luggage, or hats can alter the silhouette and features used for identification.

2) Constraints of the shooting environment

a) Low-resolution images: Surveillance cameras often produce low-resolution images, making it challenging to capture fine-grained features necessary for accurate identification.

b) Camera calibration differences: Variations in camera settings can lead to color tone mismatches and distortions, complicating cross-camera recognition.

c) Occlusions: Partial obstructions from objects, other people, or environmental factors can result in only parts of a person being visible, affecting feature extraction.

3) Dataset limitations

a) Insufficient real-world coverage: Collecting data that encompasses the vast diversity of real-world situations is difficult, leading to models that may not generalize well.

b) Privacy concerns: Legal and ethical considerations limit the ability to build large-scale datasets with varied identities, impacting the amount of available training data.

c) Lack of temporal and spatial diversity: Gathering image pairs of the same person at different times and locations is challenging, reducing the effectiveness of models in varying conditions.

4) Algorithmic challenges

a) Feature robustness: Developing algorithms that can extract features resilient to changes in posture, viewpoint, and occlusions is complex.

b) Overreliance on variable features: Many models depend heavily on easily changeable features like clothing color and hairstyle, which can lead to misidentification.

c) Long-term appearance changes: Coping with significant appearance changes over time, such as weight fluctuation or ageing, adds another layer of difficulty.

To overcome the problems, a method for person re-identification using image-to-video conversion tools is proposed in this paper. The proposed method involves matching two images taken from different viewpoints: a reference image captured in advance and a current image captured in real-time. To identify the person in concern in the current image whose image is matched to the reference image, 2D reference image is converted to 3D image and then generate images from the different aspects of the reference image. By comparing the generated images and the current image, person re-identification can be done.

The target person can be designated from a surveillance camera image as a reference. Then many images from different aspects of the target person can be generated with 2d-to-3d image conversion method. After that, the current images are acquired with the surveillance camera and are compared to the generated difference aspect images. Thus, the target person can be identified with a feature matching method between the generated image features and the current image features. This is a totally new idea of re-identification.

In the next section, related research works are to be reviewed, followed by the proposed method. Then, experiments with a small number of samples are described. After that conclusion is described together with some discussions.

II. Related Research Works

Below are some proposed methods to prevent person re-identification and their sources.

1) Data anonymization techniques: K-anonymity: A technique to ensure that each record in a dataset is indistinguishable from the records of at least k individuals.

A model for protecting privacy is discussed as L-diversity which is an extension of K-anonymity that ensures the diversity of attribute values within each group [1] together with L-diversity of privacy beyond k-anonymity [2]. T-closeness which ensures that the distribution of attribute values within each group is close to the overall distribution is discussed [3].

In terms of data masking and sampling, reduces the risk of re-identification by masking parts of the data or by random sampling is discussed and a quantitative comparison of disclosure control methods for microdata was made [4].

2) Differential privacy: A method of preventing personal identification by adding noise during data processing. Differential privacy is discussed [5].

3) Densification techniques: Protect data from unauthorized access by densifying it. Foundations of Cryptography is published from Cambridge University Press [6].

4) Privacy policies and guidelines: Establish clear policies and guidelines for collecting, using, and sharing data. European Union General Data Protection Regulation (GDPR) - Article 25 (Data protection by design and by default) is delivered [7].

5) Data aggregation and summarization: Providing aggregated or aggregated data rather than individual data prevents personal identification. Data security is well discussed in ACM Computing Surveys [8].

6) Handcrafted feature-based methods: Early ReID research used handcrafted features such as color histograms, textures, and local features to capture human features. Viewpoint invariant pedestrian recognition with an ensemble of localized features is attempted [9]. Also, person re-identification by symmetry-driven accumulation of local features is proposed [10].

7) Metric learning-based methods: A method for learning a distance metric was proposed to appropriately calculate the similarity between people. Large scale metric learning from equivalence constraints is conducted [11]. Person re-identification by local maximal occurrence representation and metric learning is proposed and tested [12].

8) Deep learning-based methods: The method of using CNN (convolutional neural network) to simultaneously learn feature extraction and classification has become mainstream. An Improved Deep Learning Architecture for Person Re-Identification is proposed [13]. Person re-identification in the wild is attempted [14]."

9) Partial feature (part-based) method: Improves recognition performance by dividing the human body into parts and utilizing information from each part. Person retrieval with refined part pooling and a strong convolutional baseline is proposed [15]. Deeply learned part-aligned representations for person re-identification are also proposed [16].

10)Attention mechanism: Improves recognition accuracy by focusing attention on important features and areas. Attention-aware compositional network for person re-identification is proposed [17]. A dual attention matching network for context-aware feature sequence-based person re-identification is also proposed [18].

11)Methods using generative models (GAN): Through data augmentation and style conversion, the amount of training data can be increased and improve the generalization performance of the model. Unlabeled samples generated by GAN improve the person re-identification baseline in vitro is proposed [19]. Person transfer GAN to bridge domain gap for person re-identification is also proposed [20].

12)Transformer-based methods: A method has been proposed to capture long-distance dependencies using a transformer architecture. TransReID: Transformer-based Object Re-Identification is proposed relatively recently [21].

13)Video generation tools: Using tools such as KLING, DeeVid.ai, Sora, etc., different viewpoint images are generated. Then the images can be used for image matching between the reference image and the current image which results in person re-identification.

14)Major datasets: Major datasets related to person re-identification are as follows:

Market-1501[1]

DukeMTMC-reID[2]

[1] https://www.kaggle.com/datasets/pengcw1/market-1501 accessed on 18 January 2025.

[2] https://www.kaggle.com/datasets/whurobin/dukemtmcreid accessed on 18 January 2025.

CUHK03[3]

MSMT17[4]

These methods are widely used to protect data privacy and reduce the risk of person re-identification. Each method has its own advantages and limitations, so it is important to choose the method that best suits particular application.

15)Evaluation methods for re-identification accuracy: Also, there are some evaluation measures for accuracy measurements of person re-identifications as follows:

mAP (mean Average Precision)

CMC (Cumulative Matching Characteristics)

Rank-1/5/10 accuracy

These related papers and datasets are referred to as follows:

CVPR (IEEE Conference on Computer Vision and Pattern Recognition)

ICCV (IEEE International Conference on Computer Vision)

ECCV (European Conference on Computer Vision)

TPAMI (IEEE Transactions on Pattern Analysis and Machine Intelligence)

III. Proposed Methods and Systems

A. Process Flow of the Proposed Person Re-Identification

The process flow of the proposed person re-identification method is illustrated in Fig. 1. The method comprises several key steps designed to accurately re-identify a person of interest captured by surveillance cameras, even when images are taken from different viewpoints.

Step 1: Detection of the Person of Interest

Reference Image Acquisition: Initially, a person of interest is selected from images acquired by surveillance cameras based on YOLO11. This image serves as the reference image for subsequent matching processes. YOLO11 provides precise bounding boxes, adjusting the center location, size, and inclination angle to accurately localize the person.

Step 2: Generation of Multi-Angle Synthetic Images

Rotated Image Generation: From each 3D model (reference image), a set of synthetic images by rotating the model at fixed intervals is generated. Specifically, 36 images are generated for each model, rotating the model by 10 degrees per image to cover a full 360-degree rotation.

Step 4: Image Matching and Re-identification

Feature Matching: Using feature matching method, the best matched person can be selected through comparison between the generated images from the different aspects and the current image.

Cross-Checking: To enhance reliability, a cross-check can be done by finding the image with the smallest difference between the current image and the rotated images generated from its own 3D model. This step validates the consistency of the current image's representation.

Optimization for computational efficiency is performed as follows:

Utilization of 2D Image Differences: To reduce computational resources, differences directly between the acquired 2D images and the generated synthetic 2D images are computed, instead of comparing complex 3D models or higher-dimensional feature representations. This approach simplifies computations while maintaining accuracy.

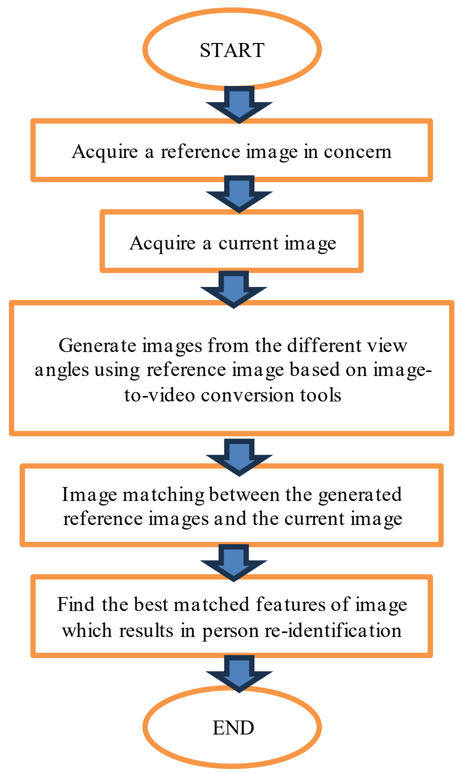


Fig. 1. Process flow of the proposed person re-identification.

Originality and advantages of the proposed method are as follows:

The key innovation of our person re-identification method lies in the integration of 3D reconstruction and multi-angle synthetic image generation to address appearance variations due to different viewpoints. By converting 2D images into 3D models and generating images from multiple angles, this method can effectively match images taken from different viewpoints, overcoming one of the primary challenges in person re-identification.

Addressing Appearance Variations: This method compensates for changes in appearance caused by different shooting angles, lighting conditions, and partial occlusions.

[3] https://www.kaggle.com/datasets/priyanagda/cuhk03 accessed on 18 January 2025.

[4] https://paperswithcode.com/dataset/msmt17 accessed on 18 January 2025.

Reducing Ambiguity Among Similar Individuals: By analyzing multi-angle representations, the method enhances discrimination between individuals who may appear similar in traditional 2D images.

Efficiency in Real-World Applications: The approach balances computational efficiency with accuracy, making it suitable for practical applications in surveillance and security systems.

B. Similarity Measure

Another key issue here is how to evaluate similarity between the 2D images in concern and the 2D images converted from the image-to-video conversion method. The method for evaluating the similarity between images should be selected taking into consideration the purpose, image features, calculation costs, etc. Here, it will be introduced some representative methods along with references.

1) Comparison at the pixel level:

a) Euclidean distance: The RGB values of each pixel are treated as vectors, and the Euclidean distance between two image vectors is calculated. The smaller the distance, the higher the similarity [22].

- Advantages: Simple and fast calculation.
- Disadvantages: Weak to translation, rotation, and scale changes of images.

b) Manhattan distance: Calculates the sum of the absolute value of the difference between the RGB values of each pixel. As with Euclidean distance, the smaller the distance, the higher the similarity [23].

- Advantages: Less susceptible to outliers than Euclidean distance.
- Disadvantages: As with Euclidean distance, it is weak to translate, rotation, and scale changes of images.

c) Histogram comparison: This method compares the brightness of histograms of images. The similarity of images can be evaluated by calculating the distance between histograms [24].

- Advantages: Robust to changes in image brightness.
- Disadvantages: Since it does not consider the content or shape of the image, histograms may be similar even for different images.

2) Feature-based comparison

a) SIFT (Scale-Invariant Feature Transform)[5]: Extracts features that are invariant to scale and rotation and evaluates the similarity between images by matching.

b) SURF (Speeded Up Robust Features)[6]: A method that can extract features faster than SIFT [25].

c) ORB (Oriented FAST and Rotated BRIEF)[7]: A feature extraction method that is even faster than SIFT and SURF [26].

d) CNN (Convolutional Neural Network)[8]: Extracts features from images using deep learning and evaluates the similarity between images by comparing them [27].

- Advantages: Highly accurate similarity evaluation is possible.
- Disadvantages: Depends on the quality of the training data. High computational cost.

3) Structural Similarity Index Measure (SSIM)[9]

a) SSIM (Structural Similarity Index Measure): A similar index that considers the characteristics of human vision. Similarity is evaluated based on three elements: image brightness, contrast, and structure [28].

- Advantages: Evaluation close to human vision is possible.
- Disadvantages: High computational cost.

4) Other

a) Earth Mover's Distance (EMD)[10]: An index that measures the distance between two distributions and is used for comparing image histograms [29].

b) Hausdorff Distance: An index that measures the distance between two-point sets and is used for comparing image shapes [30].

IMGSIM[11] is very common for image similarity measures.

In this study, just feature matching methods of ORB and SURF [31] are used for image similarity measurements.

IV. Experiments

A. Extraction of the Person in Concern from the Acquired 2D Images

For this experiment, a subset of the Market-1501 dataset, a well-known source for person re-identification, was utilized. Fig. 2(a) displays a selection of these images used in the study. From this dataset, two specific images of the individual of interest—referred to as the reference and current images—were extracted, as shown in Fig. 2(b) and (c). The images capture both the front and side views of the subject. Also, another example of a reference and a current image are shown in Fig. 2(d) and (e). These two cases are just test examples for validation of the proposed person re-identification method.

[5] https://en.wikipedia.org/wiki/Scale-invariant_feature_transform accessed on 18 January 2025.

[6] https://en.wikipedia.org/wiki/Speeded_up_robust_features accessed on 18 January 2025.

[7] https://en.wikipedia.org/wiki/Oriented_FAST_and_rotated_BRIEF accessed on 18 January 2025.

[8] https://en.wikipedia.org/wiki/Convolutional_neural_network accessed on 18 January 2025.

[9] https://en.wikipedia.org/wiki/Structural_similarity_index_measure accessed on 18 January 2025.

[10] https://en.wikipedia.org/wiki/Earth_mover%27s_distance accessed on 18 January 2025.

[11] https://github.com/totogot/ImageSimilarity/blob/main/src/ImgSim/image_similarity.py accessed on 18 January 2025.

(a) Portion of images in the Market-1501

(b) Reference image (Case1) (c) Current image of person in concern (Case1)

(d) Another reference image (Case 2) (e) Another current image (Case 2)

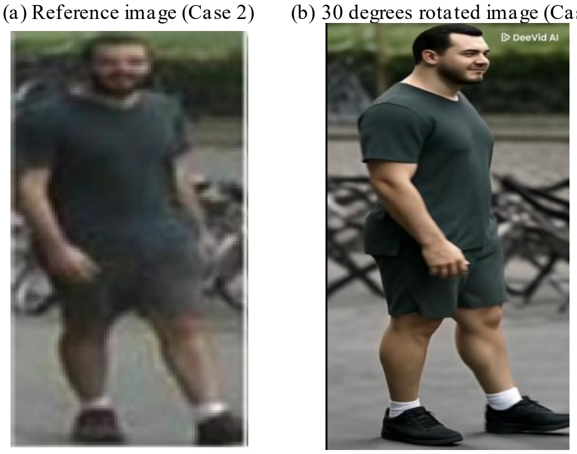
Fig. 2. Portion of Market-1501 and the extracted person in concern from the image dataset.

B. Image-to-Video Conversion Using KLING as well as DeeVid.ai

Just two examples of the resultant images obtained from the image-to-video conversion tools of KLING and DeeVid.ai are shown in Fig. 3. For these cases, the resultant images are 30 degrees rotated in horizontal direction from the original reference image.

The spatial resolution of the resultant images is superior to the reference images. It seems that the image-to-video conversion tools make models of targets and create the rotated images. Therefore, it is possible to create any rotated target image with arbitrary rotation angles. Thus, it is easier to extract features form the images for feature matching.

(a) Reference image (Case 2) (b) 30 degrees rotated image (Case 2)



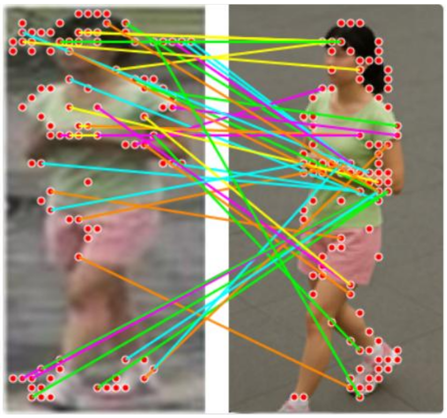
(c) Reference image (Case 1) (d) 30-degree rotated image (Case 1)

Fig. 3. Examples of the resultant images obtained from the image-to-video conversion tools of KLING and DeeVid.ai

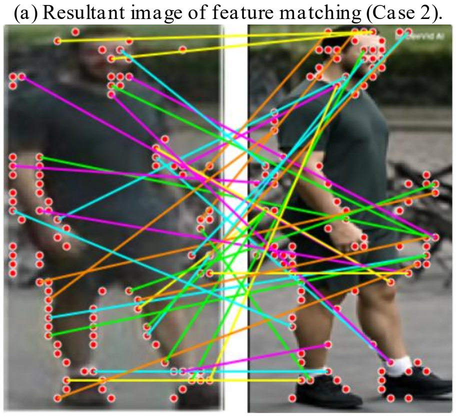
C. Feature Matching Between the Reference and the Rotated Images

ORB (Oriented FAST and Rotated BRIEF) feature point detection and matching are performed between two images. SURF is also used for confirmation of feature matching performance (Fig. 4). Feature points are detected in images (corners, edges, texture change points, etc.) and correspondences between different images are found. Red circles indicate feature points, and green lines indicate correspondence between matched feature points.

Image feature points for the reference and the 30-degree rotated images are 100. Also, the matching points between both are 31 and 42 for Case 1 and 2, respectively. Therefore, if the threshold of matching success rate is set at 30%, then the proposed person re-identification makes 100% identification success rate.



(a) Resultant image of feature matching (Case 2).



(b) Resultant image of feature matching (Case 1).

Fig. 4. Results from the feature matching.

D. *Overall Matching Accuracy for the Market 1501 Database*

In the previous section, just two examples of the success rate of the proposed person re-identification are demonstrated with a person image as a reference, together with the rotated image by every 10 degrees using KLING as well as DeeVid.ai. The highest matching success rate for the rotation angle is adopted for the determination of the identification success rate.

The same experiments are conducted for all the images included in Market-1501 with the same threshold of a matching success rate of 30. The results show more than 80% success rate. In the latest research on person re-identification in Market-1501, methods utilizing Vision Transformer have achieved excellent recognition accuracy. This method achieves high stability and accuracy by focusing on a person's distinctive features and their interrelationships, demonstrating excellent results at Rank-1 (the highest match rate).

Compared to previous CNN-based models, Vision Transformer enables more detailed feature extraction, resulting in improved Rank-1 accuracy on the Market-1501 dataset. While the specific highest success rate (the specific value of Rank-1 accuracy) varies depending on the paper or publication, advanced methods from 2024 onward have been reported to achieve accuracies of around 80%.

As such, the success rate of person re-identification in Market-1501 is steadily improving with technological advances, and the latest Transformer-based models are currently producing excellent results.

Therefore, it can be said that the proposed method reached the highest rank in terms of person re-identification success rate.

E. *Cross-Checking Between the Current Image with Different Viewing Angles and the Reference Image, and Between the Reference Image with Different Viewing Angles with the Current Image*

The proposed method introduces a robust approach to person identification through a unique bidirectional comparison system. This process consists of two complementary analysis stages:

1) Primary analysis: The system first analyzes multiple viewpoints of the current subject image, comparing each against the reference image to identify the closest match. However, this unidirectional comparison alone may not provide sufficient accuracy for reliable identification.

2) Enhanced cross-checking: To strengthen the identification accuracy, the system implements a novel bidirectional verification process:

a) Current-to-reference: Multiple viewpoints of the current subject are compared against the reference image

b) Reference-to-current: Multiple viewpoints of the reference image are compared against the current subject

This innovative bidirectional cross-checking methodology significantly enhances the reliability of person identification by:

Reducing false positives through dual verification.

Compensating for variations in viewing angles.

Providing a more comprehensive similarity assessment.

This unique approach to cross-validation represents a significant advancement over traditional unidirectional comparison methods, offering a more robust and reliable framework for person identification.

V. Conclusion

A novel approach to person re-identification using image-to-video generation is proposed.

This research presents an innovative method for person re-identification that leverages image-to-video conversion technology to address fundamental challenges in privacy protection and data security. Person re-identification—the process of identifying individuals within anonymized datasets—has historically faced significant accuracy limitations.

Traditional person re-identification methods struggle with two primary challenges:

- Appearance Variations: Subject features can vary significantly based on viewing angle and environmental conditions.

- Similar Appearances: Multiple individuals in datasets may share similar visual characteristics, complicating unique identification.

This approach significantly improves identification accuracy by:

- Compensating for viewing angle variations.
- Providing more robust feature comparison.
- Enabling better distinction between visually similar individuals.

Experimental validation demonstrates the effectiveness of this image-to-video conversion-based approach, showing marked improvements in identification accuracy compared to traditional methods. The ability to generate and analyze multiple viewpoints from 3D models provides a more comprehensive and reliable basis for person re-identification.

This improved description better organizes the information, clarifies the technical concepts, and emphasizes the innovative aspects of the proposed method while maintaining all essential information from the original text.

The detailed conclusions are as follows:

Experiments conducted using this method have shown significant improvements in person re-identification accuracy.

Effectively Addressing Appearance Variations: The use of 3D models and synthetic views accounts for changes in viewpoint, posture, and partial occlusions.

Enhancing Robustness: image-to-video creation provides a more stable representation of the individual's features, less affected by transient factors like lighting and minor attire changes.

Reducing False Matches: Improved discrimination between individuals helps in environments with many people who have similar appearances.

The proposed method offers a promising solution to the challenges faced in person re-identification by:

Mitigating Appearance Variations: Addressing one of the core issues affecting recognition rates through multi-angle analysis.

Leveraging Advanced Technologies: Utilizing cutting-edge image-to-video conversion techniques and depth estimation models enhances the robustness of the system.

Improving Practical Applicability: The method's increased accuracy brings it closer to stable use in real-world environments, benefiting applications in surveillance, security, and beyond.

VI. Future Research Works

Although it is confirmed that the proposed method does work for person re-identification with a well performance, it must be verified with the well-known databases of Market-1501, DukeMTMC-reID, CUHK03, and MSMT17. Larger-scale experiments and comparisons with state-of-the-art models (e.g. Transformer-based methods) are necessary. Also, performance comparison to the existing methods has to be made.

Continued research and development in this area can further refine the approach, potentially integrating it with other emerging technologies to overcome remaining challenges in person re-identification.

References

[1] Sweeney, L. (2002). "k-anonymity: A model for protecting privacy." International Journal on Uncertainty, Fuzziness and Knowledge-Based Systems, 10(5), 557-570, 2002.

[2] Machanavajjhala, A., Kifer, D., Gehrke, J., & Venkitasubramaniam, M. (2006). "L-diversity: Privacy beyond k-anonymity." ACM Transactions on Knowledge Discovery from Data (TKDD), 1(1), 1-52, 2006.

[3] Li, N., & Li, T. (2007). "t-closeness: Privacy beyond k-anonymity and l-diversity." In Proceedings of the 23rd International Conference on Data Engineering (ICDE '07), 106-115, 2007.

[4] Domingo-Ferrer, J., & Torra, V. (2002). "A quantitative comparison of disclosure control methods for microdata." In Confidentiality, Disclosure and Data Access: Theory and Practical Applications for Statistical Agencies (pp. 113-134), 2002.

[5] Dwork, C. (2006). "Differential privacy." In Proceedings of the 33rd International Colloquium on Automata, Languages and Programming (ICALP '06), 1-12, 2006.

[6] Goldreich, O. (2004). Foundations of Cryptography: Volume II Basic Applications. Cambridge University Press, 2004.

[7] European Union General Data Protection Regulation, Establish clear policies and guidelines for collecting, using, and sharing data. GDPR - Article 25 (Data protection by design and by default) 2024, https://gdprhub.eu/Article_25_GDPR referred on 16 December 2024.

[8] Denning, D. E., & Denning, P. J. (1979). "Data security." ACM Computing Surveys, 11(3), 227-249, 1979.

[9] Gray and Tao (2008), "Viewpoint invariant pedestrian recognition with an ensemble of localized features" European Conference on Computer Vision (ECCV), pp. 262–275, 2008.

[10] Farenzena, M., Bazzani, L., Perina, A., Murino, V., & Cristani, M. (2010), "Person Re-identification by Symmetry-Driven Accumulation of Local Features", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2360–2367, 2010.

[11] Köstinger et al. (2012), "Large Scale Metric Learning from Equivalence Constraints" IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2288–2295, 2012.

[12] Liao, S., Hu, Y., Zhu, X., & Li, S. Z. (2015), "Person Re-identification by Local Maximal Occurrence Representation and Metric Learning", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2197–2206, 2015.

[13] Ahmed, E., Jones, M., & Marks, T. K. (2015), "An Improved Deep Learning Architecture for Person Re-Identification", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3908–3916, 2015.

[14] Zheng, Z., Zheng, L., & Yang, Y. (2016), "Person Re-identification in the Wild", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3346–3355, 2016.

[15] Sun, Y., Zheng, L., Yang, Y., Tian, Q., & Wang, S. (2018), "Beyond Part Models: Person Retrieval with Refined Part Pooling (and a Strong Convolutional Baseline)", European Conference on Computer Vision (ECCV), pp. 480–496, 2018.

[16] Zhao, H., Tian, M., Sun, S., Shao, J., Yan, J., Yi, S., Wang, X., & Tang, X. (2017), "Deeply-Learned Part-Aligned Representations for Person Re-Identification", IEEE International Conference on Computer Vision (ICCV), pp. 3219–3228, 2017.

[17] Xu, J., Zhao, R., Zhu, F., Wang, H., & Ouyang, W. (2018), "Attention-Aware Compositional Network for Person Re-Identification", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2119–2128, 2018.

[18] Si, J., Zhang, H., Li, C., Zhao, X., & Tian, Q. (2018), "Dual Attention Matching Network for Context-Aware Feature Sequence based Person Re-Identification", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5363–5372, 2018.

[19] Zheng, Z., Zheng, L., & Yang, Y. (2017), "Unlabeled Samples Generated by GAN Improve the Person Re-Identification Baseline in vitro", IEEE International Conference on Computer Vision (ICCV), pp. 3754–3762, 2017.

[20] Wei, L., Zhang, S., Gao, W., & Tian, Q. (2018), "Person Transfer GAN to Bridge Domain Gap for Person Re-Identification", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 79–88, 2018.

[21] He, S., Luo, H., Wang, P., Wang, F., Li, H., & Jiang, W. (2021), "TransReID: Transformer-based Object Re-Identification", IEEE International Conference on Computer Vision (ICCV), pp. 15013–15022, 2021.

[22] Gonzalez, R. C., & Woods, R. E. (2002). Digital Image Processing (2nd ed.). Prentice Hall, 2002.

[23] Bradski, G., & Kaehler, A. (2008). Learning OpenCV: Computer Vision with the OpenCV Library. O'Reilly Media, Inc., 2008.

[24] Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision, 60(2), 91–110, 2004.

[25] Bay, H., Tuytelaars, T., & Van Gool, L. (2006). Surf: Speeded up robust features. In European Conference on Computer Vision (pp. 404-417). Springer, Berlin, Heidelberg, 2006.

[26] Rublee, E., Rabaud, V., Konolige, K., & Bradski, G. (2011). ORB: An efficient alternative to SIFT or SURF. In Proceedings of the IEEE International Conference on Computer Vision (pp. 2564-2571), 2011.

[27] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems (pp. 1097-1105) 2011.

[28] Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing, 13(4), 600-612, 2004.

[29] Rubner, Y., Tomasi, C., & Guibas, L. J. (2000). The Earth Mover's Distance as a Metric for Image Retrieval. International Journal of Computer Vision, 40(2), 99–121, 2000.

[30] Huttenlocher, D. P., Klanderman, G. A., & Rucklidge, W. J. (1993). Comparing images using the Hausdorff distance. IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(9), 850-863, 1993.

[31] Kohei Arai, Indra Nugraha Abdullah, Kensuke Kubo, Katsumi Sugawa, Methods for wild pig identifications from moving picture and discrimination of female wild pigs based on feature matching method, International Journal of Advanced Research on Artificial Intelligence, 4, 7, 41-46, 2015.

AUTHOR'S PROFILE

Kohei Arai, received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978 was also with the National Space Development Agency of Japan from January 1979 to March 1990. From 1985 to 1987, he was with the Canada Centre for Remote Sensing as a Post-Doctoral Fellow of the National Science and Engineering Research Council of Canada. He moved to Saga University as a Professor in the Department of Information Science in April 1990. He was a counselor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology from 1998 to 2000. He was a councilor of Saga University in 2002 and 2003. He was also an executive councilor for the Remote Sensing Society of Japan from 2003 to 2005. He has been a Science Council of Japan Special Member since 2012. He is an Adjunct Professor at Brawijaya University. He is also an Award Committee member of ICSU/COSPAR. He is also an adjunct professor of Nishi-Kyushu University and the Kurume Institute of Technology Applied AI Research Laboratory. He wrote 129 books and published 748 journal papers as well as 569 conference papers. He received 98 awards, including the ICSU/COSPAR Vikram Sarabhai Medal in 2016, and the Science Award of the Ministry of Mister of Education of Japan in 2015. He is now Editor-in-Chief of IJACSA and IJISA. http://teagis.ip.is.saga-u.ac.jp/index.html