International Journal of Research in Engineering, Science and Management
Volume-3, Issue-6, June-2020
www.ijresm.com | ISSN (Online): 2581-5792

Moving Object Re-identification from True-Color and Grayscale Videos

Vyshna Radh[1*], Smitha Suresh[2]
[1]*Student, Department of Computer Science and Engineering, Sree Narayana Gurukulam College of Engineering, Ernakulam, India*
[2]*Professor, Department of Computer Science and Engineering, Sree Narayana Gurukulam College of Engineering, Ernakulam, India*
*Corresponding author: vyshnaradh06@gmail.com

***Abstract*: Surveillance is the monitoring of behavior or activities of a person or an object. It can be done by using electronic equipment such as closed-circuit television (CCTV). Re-identification is an important for surveillance system. Video surveillance is one of the most active and relevant research topics in computer vision. It is well suited for a broad range of applications such as an efficient management of traffic, public safety, and military security. Moving object detection and classification are important and challenging tasks in many computer vision systems. The main application areas of object re-identification are person re-identification and vehicle re-identification. Person re-identification (Re-ID) focus on matching persons captured by different camera views. Most existing person re-identification methods focus on matching pedestrians captured by different true-color cameras. The captured pedestrian videos may be grayscale in some cases for saving storage space. In these cases, the person re-identification between true color and grayscale pedestrian videos will be needed. The proposed system is focused on moving object re-identification from true color and grayscale videos. And this is the first method for matching objects from true-color and grayscale videos. In the proposed system an efficient existing Oriented FAST and rotated BRIEF algorithm (ORB) is used for object detection. The image similarity check will be done by using Brute-Force matching algorithm.**

***Keywords*: Visual surveillance, Person re-identification, ORB, Brute-Force matching algorithm.**

1. Introduction

Surveillance is the monitoring of behavior or activities of a person or an object. It can be done by using electronic equipment such as closed-circuit television (CCTV). Video surveillance or visual surveillance is commonly used for object detection and human identification.

In a surveillance camera without overlapping vision, a recognized object is identified again after imaging conditions (including lighting conditions, object pose etc.) change, which is called object re-identification (Object Re-ID). Re-identification is the process by which anonymous personal information is matched with its true owner. Re-identification is an important capability for surveillance system as well as human computer interaction system. The main application of object Re-ID are person Re-ID and vehicle Re-ID.

Existing person re-identification problem focuses on matching persons captured between different true-color and grayscale videos. The proposed system is moving object re-identification from true-color and grayscale videos. This method focuses on matching moving objects captured by true-color and grayscale videos. In some cases the videos captured by camera are grayscale. For example, one may set cameras to grayscale mode for the purpose of saving storage space (Each pixel in the gray scale image is stored with 8 bits while each pixel in the true-color image requires 24 bits).

The processing of the presented re-identification system includes various stages. The different stages of this system is clearly explained in Fig. 1.

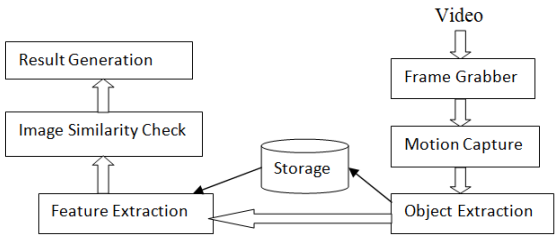


Fig. 1. Proposed system design

This paper is organized as follows. In section 2, literature survey is described. Section 3 presents proposed methodology. Finally, conclusion is delineated in section 4.

2. Literature Survey

Many researchers have introduced various techniques for person re-identification. Fei Ma et al. [1] proposed a semi-coupled dictionary pair learning (SDPL) approach. This is the first method for matching person captured between true color and grayscale video. For re-identification method, we want two sets that are probe set and gallery set. First of all, we want to calculate the distance between probe set and gallery set. After that sort the distance in ascending order. The shortest distance is the best matching pair. Here they used a new dataset called

CGVID (color to grayscale video person re-identification). They used two cameras for capturing the videos of the pedestrians. Here camera A is captured by true-color video and camera B is captured by grayscale video. First we apply feature set extraction. Here features used are corners etc. Then we apply asymmetric within-video projection matrices. Asymmetric within-video projection matrices are used to reduce the variations that are caused due to the occlusions. Then SDPL updates dictionary learning and semi-coupled mapping matrices. The semi-coupled mapping matrices will reduce the gap between features of true-color and grayscale videos.

Y. Yang et al. [2] proposed a salient color names based color descriptor (SCNCD) to describe colors. Person re-identification can be classified into two stages: feature representation and person matching [2]. This paper mainly focuses on feature representation. Color and texture are the two features used here. In image, the RGB value is mapped to the color name. The mapping function used here is probability distribution. Higher probability will be assigned to the color name which is nearer to the color [2].

Yiming Wu et al. [3] proposed an adaptive graph representation learning for video person Re-ID, which enables the contextual interactions between the relevant regional features. The advantage for this method is it is consistent among different temporal resolution for same identities. The disadvantage of this method is that it exploits pose alignment connection and feature affinity connection.

Mang Ye et al. [4] proposed a ranking aggregation algorithm. Existing methods focuses only on the similarities between probe image and gallery image. And it doesn't consider the dissimilarities between them. But this method focuses on both similarities and dissimilarities. The main idea behind this method is that it contains an original ranking list. The original ranking list consists of probe image, strongly similar galleries, neutral galleries, strongly dissimilar galleries. If the probe images matches with the strongly similar galleries, then it will be pulled. If the probe images matches with the strongly dissimilar galleries, then it will be pushed.

Ancong Wu et al. [5] addressed the RGB-IR cross-modality Re-identification problem and contribute a dataset named SYSU-MM01, including RGB and IR images. This is the only method for matching RGB and IR images. The disadvantage of this method is that it is a challenging task due to large cross modality variations between RGB and IR images.

Liang Zheng et al. [6] introduced a video re-identification dataset called Motion Analysis and Re-identification Set (MARS). It is the video extension of the Market- 1501 dataset. MARS is the largest video re-id dataset. The pedestrians are captured by using six cameras. It consists of high quality images. For detecting the pedestrians deformable part model is used. GMMCP tracker is used to fill the missing result. The disadvantage is that it is designed for matching between truecolor videos.

3. Proposed Methodology

The moving object re-identification from true-color and grayscale videos focuses on matching moving objects captured from true-color and grayscale videos. Moving object re-id includes three stages: Object Extraction, Feature Extraction and Image similarity check.

The first step in the visual surveillance or video analysis is the moving object extraction. After extracting the object, feature extraction is done. For feature extraction, Oriented FAST and rotated BRIEF (ORB) algorithm is used. The final step is to check the similarities of objects between true-color and grayscale videos. The image similarity check will be done by using Brute-Force matching algorithm.

In the proposed system, the first step is to convert the captured videos into successive frames. These frames are used for further processing. Then the next step is to capture the motion. Motion capture is the process of recording the moving objects or people. Then the next step is object extraction. The object extraction is done by using contours. After extracting the object, feature extraction is done. For feature extraction, Oriented FAST and rotated BRIEF (ORB) algorithm is used. The final step is to check the similarities of objects between true-color and grayscale videos. The image similarity check will be done by using Brute-Force matching algorithm.

A. Object Extraction

In the block diagram we have seen that the primary step of the video object classification system is to capture the video and convert this video into frames. After converting it into frames we can move on to the moving object detection step. In this process the object is extracted based on contour object extraction from images.

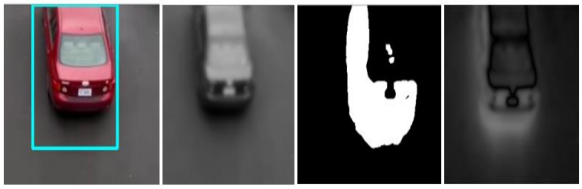
Contour is a curve joining all the continuous points (along the boundary), having some color or intensity. The contours are a useful tool for object detection and recognition. For better accuracy, use binary images. So before finding contours, apply threshold.

The first step is capture the true-color and grayscale videos. After that convert these videos into frames. The next step is to find the difference between two frames. To find the difference between two frames, we use Gaussian blur filter. Gaussian blur filter is used to blur images and remove noise. For finding the difference, we have given a threshold value of 50. After finding the difference, the images will be in gray. Then that gray image will be converted into black and white image. The pixel value below 50 will be converted to black. And pixel value above 50 will be converted to white.

Binary images are images whose pixels have only two intensity values. They are displayed as black and white. Numerically, the two values are 0 for black, and either 1 or 255 for white. Binary images are produced by thresholding a grayscale or color image, in order to separate an object in the image from the background. The color of the object (usually white) is referred to as foreground color. The rest (usually

black) is referred as background color. However, depending on the image to be thresholded, this polarity might be inverted.

Then the next step is to apply dilation filter. If there are any issues in edges, then to normalize the edges, dilation filter is used. Dilation filter increases the object area. After applying dilation filter, we will get a pure black and white image. Then we will find contours from that image. If it is a black and white image, then there will be a transition from one color to another. The curve joining all the pixels from that transition is called contours. Finding contours is like finding white object from black background. The changed portion will become white. And the portion without change will become black. If we cut the white portion from the image, we will get the changed portion. If the same portion is cut in the original image, then we will get the object. In this way, we will extract the objects.



(a) (b) (c) (d)

Fig. 2. Moving Object Extraction for True-Color Camera (a) color frame (b) gray frame (c) threshold frame (d) delta frame

Fig. 2 shows moving object extraction for true-color camera. Here Gray frame is obtained by converting color image to gray mode. Delta frame is obtained by calculating the absolute difference between the first frame and the gray frame. In threshold frame, the changed portion will become white and the portion without change will become black.

(a) (b) ` (c) (d)

Fig. 3. Moving Object Extraction for Grayscale Video (a) color frame (b) gray frame (c) threshold frame (d) delta frame

B. Feature Extraction

The core of object Re-ID algorithm is feature extraction model; the effectiveness of the whole algorithm is almost determined by this part. In other words, the essence of Re-ID is to compare the similarity or distance between the features extracted from two images. Image features mainly include color, texture, shape and spatial relationship feature. Feature extraction refers to the use of computer to extract image information to determine whether each image pixel belongs to an image feature. Features are the best way to describe patterns, and we often think that each dimension of a feature can describe a pattern from a different perspective.

The key portion of feature based system is to consider what strong and robust feature will be extracted. Feature correspondence is also the challenging problem because a feature point in one region of image may have many similar points in another region of this image or another image and it tends to accuracy in feature correspondence uncertainty. Selecting the right and covered feature plays a very important role in video object classification. The selected features are combined into a single feature vector (FV) for classifying the object. There are many features based on texture, color, gradient, and statistics available for object classification.

For feature extraction, we use Oriented FAST and rotated BRIEF algorithm. ORB is a combination of FAST keypoint detector and BRIEF descriptor with some added features to improve the performance. FAST is Features from Accelerated Segment Test used to detect features from the provided image. It also uses a pyramid to produce multiscale-features. Now it doesn't compute the orientation and descriptors for the features, so this is where BRIEF comes in the role.

ORB uses BRIEF descriptors but as the BRIEF performs poorly with rotation. So ORB does is to rotate the BRIEF according to the orientation of keypoints. Using the orientation of the patch, its rotation matrix is found and rotates the BRIEF to get the rotated version. ORB is an efficient alternative to SIFT or SURF algorithms used for feature extraction, in computation cost, matching performance. ORB extracts keypoints and descriptor.

A keypoint and a descriptor define a feature. A keypoint (or interest point) is defined by some particular image intensities "around" it, such as a corner. A keypoint is used for deriving a descriptor. Not every keypoint detector has its particular way for defining a descriptor. A descriptor is a finite vector which summarizes properties for the keypoint. A descriptor is used for classifying the keypoint. ORB has an absolute advantage in speed. The greatest feature of this algorithm is fast and having rotational invariance and reducing sensitivity to noise.

C. Image Similarity Check

Image similarity check is used to check the similarity of the selected true color image in a grayscale image. For image similarity check, we will use Brute force matching algorithm.

After extracting the keypoints and descriptors by using ORB, Brute force matching algorithm will be done to find the similarities of the features. It takes the descriptor of one feature in first set and is matched with all other feature in second set using normal hamming distance calculation. And the closest one is returned. The shortest distance is the truly matching pair.

Fig. 4. Probe Image

Fig. 4 shows the probe image. The probe image is from true-color video. Then identify the probe image from grayscale video. Fig 5 shows re-identification of true-color image from grayscale video. It will show the exact position where the image is in grayscale video.

Fig. 5. Re-identification of True-Color Image from Grayscale Video

4. Conclusion

To study more on person re-identification, a literature survey was done. From the survey, it is clear that each method has limitations and still need to work on person re-identification. The proposed method is moving object re-identification from true color and grayscale videos. Moving object re-identification is used to identify and match the moving objects captured between true-color and grayscale videos. The advantage of using grayscale videos is that it saves storage space and less information is needed to store a pixel. In proposed method, contour-based object extraction is used to extract the objects. For feature extraction, Oriented FAST and Rotated BRIEF algorithm (ORB) is used. ORB algorithm has a strong real-time performance and has a great improvement in computing speed. The greatest feature of this algorithm is fast and having rotational invariance and reducing sensitivity to noise. The final step is image similarity check. It is used for re-identification process. For image similarity check, Brute force matching algorithm is used.

References

[1] F. Ma, X.Y. Jing, X. Zhu, Z. Tang and Z. Peng, "True-Color and Grayscale Video Person Re-identification," IEEE, US, 2019.
[2] Yang, J. Yan, S. Liao, D. Yi and S. Z. Li, "Salient Color Names for Person Re-identification" European Conference on Computer Vision (ECCV), Switzerland, pp. 536–551, 2014.
[3] Y. Wu, O.E.F. Bourahla, X. Li, F. Wu, and Q. Tian, "Adaptive Graph Representation Learning for Video Person Re-identification," IEEE, US, 2018.
[4] M. Ye, C. Liang, Y. Yu, Z. Wang, Q. Leng, C. Xiao, J. Chen and R. Hu, "Person Re-identification via Ranking Aggregation of Similarity Pulling and Dissimilarity Pushing," IEEE, US, 2016.
[5] A. Wu, W.-S. Zheng, H. Yu, S. Gong and J. Lai, "RGB-Infrared Cross Modality Person Re-identification," IEEE 2017.
[6] L. Zheng, Z. Bie, Y. Sun, J. Wang, C. Su, S. Wang and Q. Tian, "MARS: A Video Benchmark for Large-Scale Person Re-identification," European Conference on Computer Vision (ECCV), Switzerland, 2016.
[7] Ethan Rublee, Vincent Rabaud, Kurt Konolige and Gary Bradski, ORB: an efficient alternative to SIFT or SURF, IEEE Inter. Conf. on Comput. Vis. (ICCV), 2011.