Metrol. Meas. Syst., Vol. 29 (2022) No. 1, pp. 53–74
DOI: 10.24425/mms.2022.138548

PAN POLSKA AKADEMIA NAUK

METROLOGY AND MEASUREMENT SYSTEMS
Index 330930, ISSN 0860-8229
www.metrology.wat.edu.pl

M&MS

RFID TAG GROUP RECOGNITION BASED ON MOTION BLUR ESTIMATION AND YOLOv2 IMPROVED WITH THE GAUSSIAN ALGORITHM

Lin Li[1,2)], Xiao-Lei Yu[1,2)], Zhen-Lu Liu[1)], Zhi-Min Zhao[1)], Ke Zhang[1)], Shan-Hao Zhou[1)]

1) *College of Science, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China,* (*lilin@nuaa.edu.cn*, +55 55 99608 1406, ✉ *nuaaxiaoleiyu@126.com*, *550138109@qq.com*, *zhaozhimin@nuaa.edu.cn*, *2055282625@qq.com*, *970270047@qq.com*)
2) *National Quality Supervision and Testing Center for RFID Product Jiangsu, Nanjing 210029, China*

Abstract

Effective recognition of tags in the dynamic measurement system would significantly improve the reading performance of the tag group, but the blurred outline and appearance of tag images captured in motion seriously limit the effectiveness of the existing tag group recognition. Thus, this paper proposes passive tag group recognition in the dynamic environment based on motion blur estimation and improved YOLOv2. Firstly, blur angles are estimated with a Gabor filter, and blur lengths are estimated through nonlinear modelling of a Generalized Regression Neural Network (GRNN). Secondly, tag recognition based on YOLOv2 improved by a Gaussian algorithm is proposed. The features of the tag group are analyzed by the Gaussian algorithm, the region of interest of the dynamic tag is effectively framed, and the tag foreground is extracted; Secondly, the data set of tag groups are trained by the end-to-end YOLOv2 algorithm for secondary screening and recognition, and finally the specific locations of tags are framed to meet the effective identification of tag groups in different scenes. A considerable number of experiments illustrate that the fusion algorithm can significantly improve recognition accuracy. Combined with the reading distance, the research presented in this paper can more accurately optimize the three-dimensional structure of the tag group, improve the reading performance of the tag group, and avoid the interference and collision of tags in the communication channel. Compared with the previous template matching algorithm, the tag group recognition ability put forward in this paper is improved by at least 13.9%, and its reading performance is improved by at least 6.2% as shown in many experiments.

Keywords: RFID, YOLOv2, neural network, GRNN.

1. Introduction

With rapid development of the *Internet of Things* (IoT), the supply chain management has promoted the development of non-contact technology. *Radio Frequency Identification* (RFID) is one of the non-contact technologies widely used in access control, parking lots, patient tracking,

Article history: received June 15, 2021; revised September 24, 2021; accepted October 21, 2021; available online November 10, 2021.

toll gate payment systems, defense applications, production line automation, logistics tracking, *etc.* [1,2]. To get rid of the limitations of inventory both management and industrial process, RFID aims to improve the speed and accuracy of the system through data management and collection.

When locating objects with RFID, collisions between tags and antennas inevitably occur in the channel, which seriously affects the tag positioning problem [3]. To calculate *Received Signal Strength* (RSS), the power is calculated at the reader or the base station. RSS measurements require at least three readers to analyze. In [4], tag positions are most likely to be estimated via transmission level or RSS by stationary receivers, but RSS measurement is not reliable owing to the change of the environment. The LANDMARC system finds the closest target tag by the RSS value in [5]. Vire improves positioning accuracy by virtual marking [6]. In [7], high-precision estimation of position is improved by reducing the number of target adjacent labels. In [8], RFID tags are roughly located by rotating antennas, and more positioning information can be acquired through the dynamic environment. Then, the distance of relative incident angle between the tag and the antenna is acquired through *Received Signal Strength Indication* (RSSI). These methods mainly focus on improving reader accuracy, RSS, and the positioning method. Yet, the location of tags is inevitably affected by the interference of tag communication.

To avoid channel interference, some scholars have tried to combine RFID and computer vision to improve positioning accuracy. Computer vision eliminates interference between channels by image processing and the neural network. In [9], the authors combine RFID with computer vision effectively through *Dempster-Shafer* (DS) theory, according to the prior probability error distribution. In addition, the feedback of high-confidence tracking results and RFID signal corrects the wrong vision to overcome the drift problem of target aggregation and occlusion, thus achieving monitoring of actual targets. In [10], the fusion algorithm of fine-grained positioning and tracking of marked objects can find targets on the screen accurately. In [11], the multi-path propagation model and the dual antenna solution are proposed to minimize the phase effect of multi-path interference. Meanwhile, the *Region of Interest* (ROI) of RFID tag objects is extracted by image processing. Then, the phase offset caused by multi-path interference eliminates the ROI to minimize the uncertainty of the tag's position caused periodically by the radio frequency phase. So, the fusion of RFID and computer vision may reduce multi-path interference to a certain extent. It is still hard work to improve the location information of RFID tags t computer vision.

In RFID tag identification based on computer vision, all the tags access the gate to simulate the actual goods entering and leaving the warehouse. When all the tags enter and exit the gate simultaneously, it is called a tag group. The tag groups are located through the RF system, which will make the location information inaccurate due to the interference of the surrounding environment. The images of tag groups with different geometric distributions are processed and analyzed to find the location of tags through image processing, which allows to avoid the collision and interference between tags in the channel [12–15]. The collision between tag groups is eliminated geometrically, which is called physical anti-collision. In [12,16], the knife-edge and the Wiener filter method are proposed, and then the three-dimensional position measurement of RFID tags group are predicted by a *Deep Brief Network* (DBN) which overcomes the shortcomings of a *Back Propagation* (BP) neural network, such as easily falling into local optimum and long training time due to random initialization of weight parameters. In [13], multi-label image denoising based on a fast and flexible deep convolution neural network is proposed, and multi-labels are matched by template matching. Then, the relationship between the pixel and spatial coordinates was built by photogrammetry to get 3D positions in the multi-label world coordinate system. In [14], 3D coordinates and reading distance of multi-tags are optimized and analyzed by the *Mind Evolutionary Algorithm* (MEA) to acquire optimal 3D distribution. The reading distance is predicted by a *Support Vector Machine* (SVM), thereby obtaining optimal geometric distribution

Metrol. Meas. Syst., Vol. 29 (2022), No. 1, pp. 53–74
DOI: 10.24425/mms.2022.138548

and improving the reading performance [17]. These methods eliminate channel and multi-path interference by optimizing tag image through the neural network and image processing.

To accurately and quickly read the positioning information of the tag group, the RFID tag group image is collected by a double *Charge Coupled Device* (CCD) camera. In this paper, the information in the horizontal direction of the tag group is obtained by a vertically placed camera, and the image information in the vertical direction of the tag group is collected by a horizontally placed camera. This camera placement method is called the horizontal-vertical CCD camera. There is motion blur in the acquired tag images owing to the relative motion between the CCD cameras and tags. To improve the accuracy of 3D measurement, it is necessary to remove the motion blur and restore the latent tag group images. The purpose of image restoration is to retrieve degraded images with prior knowledge. First, a model of the blurry image is established, and then latent sharp images are recovered from the degraded image.

In the restoration of the motion blur image there are two key problems to build a degradation model more accurately and precisely capture the model parameters. The *Point Spread Function* (PSF) is estimated by the characteristics of the tag group image in the parameter estimation, and then the blurry image is restored by the fringes in the frequency domain. Some scholars have studied the distribution rule of dark stripes in the spectrum of blurry images and proposed corresponding parameter detection algorithms [18]. The spacing between the lines of the blurred image spectrum is determined by the Hough transform [19]. The blur angle is calculated by the Radon transform [20]. But any small error may cause a big change in the estimation of the blur angle. The Hough transform cannot directly process grayscale images, and the trajectory of the straight line is relatively blurry [21]. When line pixels of a straight are incomplete, the accuracy is not high and the amount of calculations necessary is very large. The Radon transform has a fine adaptability to the detection of blur angle, but the estimated value is not a sufficiently accurate when the blur angle is 45. The estimated values of the Gabor filter in all directions are accurate enough, and the operation efficiency is equivalent to that of the Radon transform [22].

In addition, the 3D distribution measurement of the tag group has been studied. In [15], the 2D distribution of the multi-tag image is extracted by flood filling, and the 3D geometric distribution of the multi-tag is obtained through the *Direct Linear Transformation* (DLT). In [12], the distribution of tags was found through morphology and template matching. Yet, these methods require a large amount of calculations and long running time. When the image rotates and the size changes, it is difficult to meet the system requirements. As an end-to-end manner, the YOLO series has been widely processed and applied in object recognition [23–25]. YOLO recognizes objects by dividing the image into grid cells, rather than through the two-stage detector of the region recommendation method. The detection speed of YOLO is much bigger than that of conventional methods, but the positioning error is relatively large owing to the processing of grid elements, and the detection accuracy needs to be improved. Compared with YOLO, YOLOv2 improves the detection accuracy by batch normalization of the convolution layer, anchor box, and fine-grained features. However, when the size of the object relative to the image is very small or the object is blocked, the detection accuracy needs to be improved. To overcome the shortcomings of YOLOv2, YOLOv3 is composed of multiple convolution layers, and the accuracy is improved through a depth network. YOLOv3 selects the residual jump link to solve the problem of gradient disappearance and detects the feature images of different scales through the upsampling method, so YOLOv3 can detect objects of different sizes in the image. To further improve the accuracy and robustness of tag group recognition, this paper selects YOLOv2 improved by the Gaussian algorithm to recognize tag groups. Thus, this paper proposes a tag group recognition method based on motion blur estimation and YOLOv2 improved by the Gaussian algorithm. The main contribution of research proposed in this paper is twofold, as follows:

1. The prior knowledge formed when the original sharp tag group image is degraded, that is, the PSF, is the key to restore the motion blur image. The blur length and blur angle of the PSF are the keys to restoration, and as such directly affect the quality of image restoration. Therefore, this paper estimates the blur angle through the Gabor filter and the blur length by a *General Regression Neural Network* (GRNN), which can improve the accuracy of PSF parameter estimation. Finally, the tag group images are deblurred by the Wiener filter;
2. A tag group recognition algorithm combining Gauss and YOLOv2 is proposed. In the fusion algorithm, firstly, the tag foreground is extracted by an improved mixture Gaussian algorithm to eliminate the static interference and reduce the scope of tag recognition. Then, the data is trained and the characteristics of the tag group are extracted by the object recognition algorithm YOLOv2. Finally, the position and scope of the tags are predicted and the tags are framed. Compared with YOLOv2, the improved YOLOv2 heightens the real-time performance of prediction, reduces the false detection rate, which verifies the effectiveness and feasibility.

The rest of this paper is arranged as follows. Section 2 describes system architecture. The image deblurring principle and experiment are described in Section 3. Section 4 is the principle and experiment of tags group recognition by improved YOLOv2. The conclusions are given in Section 5.

2. System architecture

To optimize the reading performance of tag groups, this paper presents the simulation of the transmission of goods based on an RFID semi-physical verification platform. RFID tags are attached to the goods in transit. According to different power supply modes, there are three types of RFID tags: active tags, passive tags, and semi-active tags [26]. Active tags can provide energy for themselves without the assistance of the outside world. Passive tags get the energy needed by the reader in the magnetic field, which is small and easy to carry. The battery power supply in the semi-active tags only supports the circuit in the tag that requires power supply to maintain data or the voltage required by the tag chip. The passive tag is selected in this paper. When the tag enters the range of the reader, the tag extracts the required power from the RF signal emitted by the reader. Tag information is transmitted to the readers by reflection modulation, in which it would inevitably interfere with the surrounding electromagnetic information and other tags and readers, and then affect the transmission of tag information.

The structure diagram of RFID tag group recognition measurement system is shown in Fig. 1. It is mainly composed of the guide rail, readers, antennas, RFID tags, vertical and horizontal CCD cameras group, *etc*. The image acquisition devices are fixed horizontal-vertical CCD cameras group. The guide rail simulates the movement of logistics transportation, on which a turntable is mounted. Brackets are placed on the turntable, and RFID tags are clamped on the bracket. The horizontal CCD is placed perpendicularly to the direction of the conveyor belt, and its main axis is parallel to the ground, which allows it to acquire tag group image from left to right to obtain the side view of the tags. The vertical CCD is vertically downward relative to the conveyor belt, and its main axis is perpendicular to the ground, which mainly collects the top image of the tag group from top to bottom. The light of the two cameras is relatively vertical. The computer communicates with the horizontal and vertical CCD camera groups through a data transmission interface. The image becomes blur due to relative motion between the tag and the camera. Firstly, the obtained image is deblurred by motion blur estimation, and then the position of the tags is recognized in the restored sharp image. Through image processing and analysis, the coordinates

Metrol. Meas. Syst., Vol. 29 (2022), No. 1, pp. 53–74
DOI: 10.24425/mms.2022.138548

of the tags in the z direction can be obtained from the image collected by the horizontal CCD, and the coordinates of the tags in x and y directions can be calculated from the image collected by the vertical CCD.

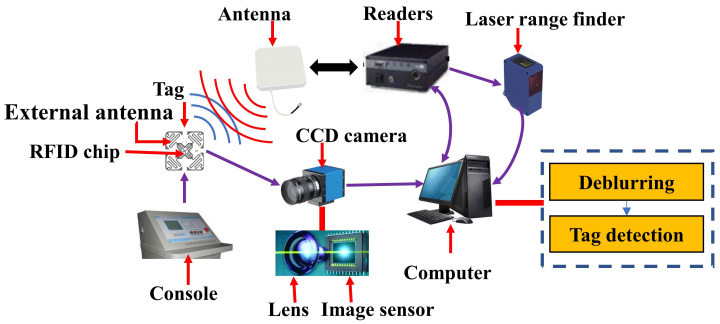


Fig. 1. Structure diagram of the RFID tag group recognition measurement system.

3. Motion blur of RFID tag group

3.1. Motion blur parameters

When acquiring tag group image, the image may be degraded due to relative motion between tags and the camera, which is motion blur. The motion degradation is simulated by a 2D linear translation invariance. The blurry image is represented by $g(x, y)$, and the original image is $f(x, y)$. $g(x, y)$ is the convolution of the original image $f(x, y)$ and the PSF. So, image deblurring is a deconvolution operation. The degradation model is shown in Fig. 2, and the mathematical equation is shown as follows:

$$g(x, y) = f(x, y) * h(x, y) + \eta(x, y), \tag{1}$$

$*$ represents 2D linear convolution, $\eta(x, y)$ is additive noise, and $h(x, y)$ is the PSF. The model expression in the frequency domain is as follows:

$$G(u, v) = F(u, v)H(u, v) + N(u, v). \tag{2}$$

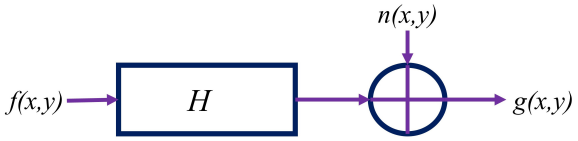


Fig. 2. The degradation model of blurry image.

Among them, $G(u, v)$, $F(u, v)$, $H(u, v)$ and $N(u, v)$ are the Fourier transform of blurry image, original image, PSF and noise, respectively. Ignoring $n(x, y)$ in (1), when $g(x, y)$ is known, $f(x, y)$ is solved, which transforms the problem into solving the PSF $h(x, y)$. The motion blur restoration is the inverse operation of the degraded model, in which the accuracy of the PSF is crucial to the restoration effect. The PSF is uncertain, and that can only be estimated by extracting degraded information from the blurred image.

For 2D linear translation-invariant system, the mathematical description of the PSF blur for the uniform linear motion is as follows:

$$h(x,y) = \begin{cases} 1/L & \sqrt{x^2+y^2} \le L/2, \quad y/x/= /\tan\theta \\ 0 & \text{in other cases} \end{cases}. \tag{3}$$

The PSF has two parameters, motion blur length L and blur angle θ, which are then transformed into the solution process of two parameters. According to (3), the frequency response of the PSF is sin c function. The blur image of the tag group and its spectrum diagram are shown in Fig. 3. The performance of blurry image restoration depends on the accuracy of PSF parameter estimation. Therefore, it is necessary to accurately estimate blur length and angle from the motion blur function. To filter blur images, the blur angle can be estimated with the Gabor filter based on its response in the frequency domain. The blur length can train the neural network to evaluate. Once the blur angle and length are acquired, the PSF is constructed, and finally, the deblurred image is restored by the Wiener filter.

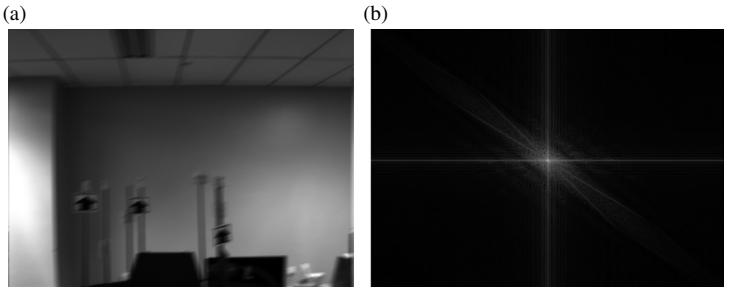


Fig. 3. Tag motion blur image, (a) origin image, (b) corresponding spectrum.

As can be seen from Fig. 3, the direction of parallel black lines and the direction of motion blur angle are perpendicular to each other. The angle of parallel lines is calculated by all line detection algorithms, that is, the angle of the motion blur. In this paper, the blurred image is filtered with the Gabor filter in the frequency domain, and the blur angle is estimated according to its response.

3.2. Motion blur parameter estimation

3.2.1. Blur angle

In this paper, the frequency response of blur images is adopted. The relationship between the blur angle and the display direction is as follows:

$$\alpha = \phi + 90. \tag{4}$$

Many line detection algorithms can detect the line direction in the blur image spectrum, such as the Hough transform or the Radon transform. The Hough transform needs a threshold to determine the points on the line. The threshold corresponding to different images is different. Any small error in threshold estimation may lead to a large deviation in blur angle estimation. To solve the problem of threshold estimation, the angle of motion blur is determined by the Gabor filter.

Metrol. Meas. Syst., Vol. 29 (2022), No. 1, pp. 53–74
DOI: 10.24425/mms.2022.138548

The Gabor transform is a windowed Fourier transform, which can extract relevant features in different scales and directions in the frequency domain. It is commonly used in image segmentation, texture recognition, and angle extraction, *etc*. The Gabor filter is a Gaussian filter modulated by a sine wave. The equation of a typical 2D Gabor filter is as follows [22]:

$$G(x,y) = \frac{1}{2\pi\sigma_x\sigma_y}\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)\right]\exp\left[-j\omega\left(x\cos\varphi + y\sin\varphi\right)\right], \tag{5}$$

where the first is the tuning function, and the second is the window function. ω is the modulation signal frequency, θ is the modulation signal angle, σ_x, σ_y are the standard deviations in the x and y directions, respectively. The response of the filter varies with the orientation parameter, so the blur angle can be calculated by the directional parameter of the 2D Gabor filter. The 2D Gabor filter is convoluted with the spectrum of the blurred image, and then the response in different directions is received by changing other parameters. The line orientation in the spectrum of the blurred image can directly affect the judgment of blur angle. The Gabor filter determines the blur angle, extracts the features of the blur image and determines the direction of the motion blur, which can effectively reduce the impact of this problem. Assuming that $I = \lg(G(u,v))$ is a grey image in the frequency domain, the response of the Gabor filter depends on the frequency and direction of the input image. When the Gabor filter detects the motion direction in the spectrum of the blur image, the blurred image is Fourier-transformed to get I. Secondly, the Gabor filter can only change the filter direction and keep other parameters unchanged. Gabor filters in different directions are convoluted with I to obtain the response of each angle $R(\varphi) = G(x,y) * I$. The angle corresponding to the maximum value is taken as the angle of the point, that is, the response at a specific frequency and direction. For each angle, the direction corresponding to the maximum value of the convolution is the angle of the point. The direction corresponding to its maximum value is calculated by the L_2 norm. So, for each φ, the L_2 norm generated by the convolution is calculated. The maximum value of L_2 norm corresponds to the blur angle θ^{θ}.

3.2.2. Blur length

Another parameter of the motion blur is the blur length which describes the distance the tag covers in the image pixels during the exposure time. To predict the blur length of a specific blurred image, the sum of amplitudes of the Fourier series corresponding to the blurred image is as input. The Fourier feature is one of the simplest features in the frequency domain. The nonlinear relationship between the sum of amplitudes of the Fourier series and the blur length can be modelled by a *Generalized Regression Neural Network* (GRNN).

The GRNN is a radial basis function neural network which has strong nonlinear mapping ability and robustness, and has great advantages in approximation ability and learning speed [27, 28]. The GRNN structure, which better solves nonlinear problems, is shown in Fig. 4. The GRNN is composed of four layers, and they are input, output, mode, and sum layer.

In training, 50 tag group images of each scene are selected as the data set, of which 40 images are as the training data sets and 10 as the testing data. There are 6 scenes in total, there are 240 training image data sets and 60 test image data sets. The range of the blur length is from 2 to 45 and the step size is 1. The SUMFC of each image was calculated with a total of 352 training samples, of which 70 are as test data. The error of the final control result is less than 0.01. Then, the blur length is estimated by the trained GRNN.

Tag group image information is collected by the CCD camera. In this paper, a series of images are blurred in different scales and directions. The blur angle is taken as $0 < \phi < 90$, and the

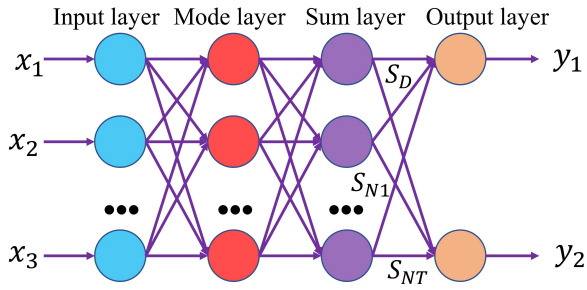


Fig. 4. GRNN structure.

direction of the blur is $0 < L < 45$. Blur angles are estimated with different Gabor templates to make them closer to actual blur angles. In the Gabor filter, the aspect ratio parameter is 0.5, the wavelength is 20 pt, the direction is 0°, the phase offset is 0°, and the bandwidth is 1. Then, Fourier coefficients of the horizontal blurred image in the frequency domain are used as the input to train the neural network and estimate the blur length. Fig. 5 shows the convergence curve of the generalized regression neural network.

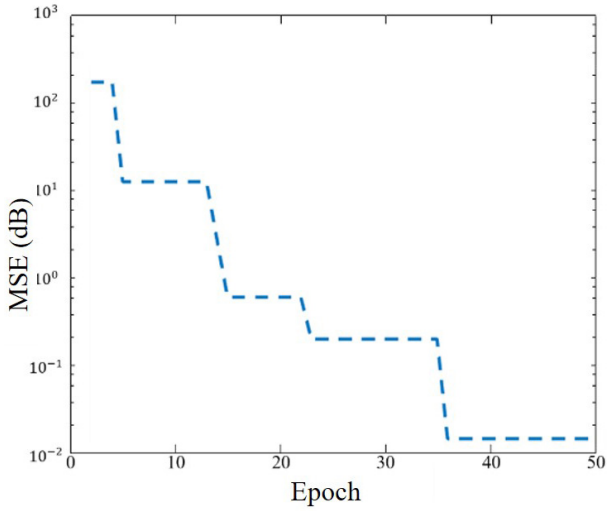


Fig. 5. Convergence curve of the GRNN.

3.3. Image motion deblurring results and analysis

To verify the effectiveness of the proposed method, the motion blur images of the tag group with different angles and lengths are analysed in this paper. The Wiener filtering follows the *Minimum Mean Square Error* (MMSE) criterion, which is as follows [29]:

$$e^2 = E\left\{\left(f - \hat{f}\right)\right\}^2. \tag{6}$$

Metrol. Meas. Syst., Vol. 29 (2022), No. 1, pp. 53–74
DOI: 10.24425/mms.2022.138548

The core idea is to minimize the mean square error between the original image and the restored image. According to this principle, the Wiener filtering equation can be deduced as follows [30]:

$$F(u,v) = \frac{1}{H(u,v)} * \frac{|H(u,v)|^2}{|H(u,v)|^2 + s\dfrac{P_n(u,v)}{P_f(u,v)}} G(u,v), \tag{7}$$

where $|H(u,v)|^2 = H^*(u,v)H(u,v)$, $H(u,v)$ is the frequency domain form of the PSF, $G(u,v)$ is the frequency-domain form of the blur image, $P_n(u,v)$ is the noise power spectrum, and $P_f(u,v)$ is the original image power spectrum. Let $K = sP_n(u,v)/P_f(u,v)$. The experimental results show that when K takes the empirical value of 0.0039, the restoration effect of the obtained image is the best.

Fig. 6 shows the static tag group image captured by the horizontal camera. By simulating the actual situation, the tag group image has different blur lengths and angles in motion. Images with different blur lengths and angles are selected as the input to the proposed estimation scheme. The Gabor filter calculates the blur angle of each image, and the GRNN calculates the blur length of each image. Table 1 shows blur angle estimation of the restored image, which shows that the estimation of each angle with the Gabor filter is very efficient. Blur length estimation of the restored image is shown in Table 2(a-d) which represents blurry images at different times.

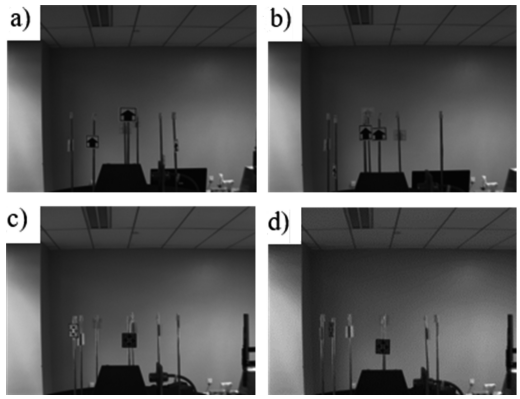


Fig. 6. The static tag group image.

Table 1. Blur angle estimation of the restored image.

Image	Angle θ	Radon	Ours
(a)	15	14	14
(b)	30	29	29
(c)	40	37	39
(d)	45	41	47
(e)	60	63	61
(f)	75	78	74
(g)	90	93	88

Table 2. Blur length estimation of the restored image.

Image	Length L	Radon	Ours
(a)	10	11	10
(b)	15	19	14
(c)	30	32	29
(d)	45	41	44

In Table 1, it can be found that the Radon transform has an ideal accuracy for blur angle estimation, but the estimation error of image blur angle at 45° is relatively large. At the same time, the Gabor filter has good angle estimation at all angles. In Table 2, one can see that the estimation of the blur length by the GRNN has obvious advantages over the traditional method, in which the error between estimated and actual blur length is smaller. The PSF is constructed by the estimated blur parameters (θL), and the degradation model is established. In this paper, the blurred image is restored with the Wiener filter, and *Mean Square Error* (MSE) is used as the evaluation index to evaluate the results of image restoration as shown in Table 3. MSE between the restored image and the original image is small, which indicates that the method in this paper can effectively remove the motion blur from the image.

Table 3. MSE of the restored image.

Image	Length L	Angle θ	Radon	Ours
(a)	10	15	3.68E-01	1.22E-02
(b)	15	30	7.69	3.26E-02
(c)	30	40	5.98E-01	2.29E-02
(d)	30	45	5.36E-01	2.32E-02

Fig. 7 is a comparison of the motion deblurring results of our research with other methods. In this paper tag group image restoration is proposed based on motion blur parameter estimation,

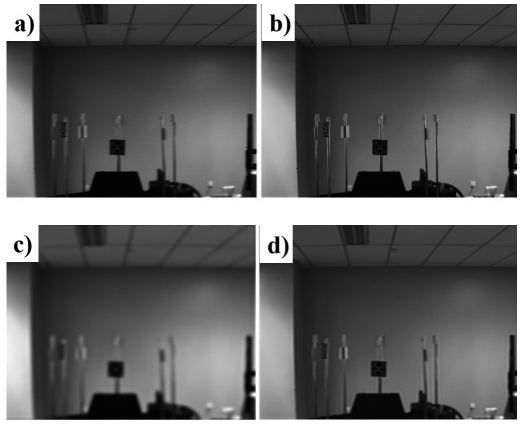


Fig. 7. Comparison of the motion deblurring results achieved in our study with other methods, a) blurred image, b) restoration results of our method, c) restoration results of constrained least square method, d) restoration results of Lagrange operator method.

Metrol. Meas. Syst., Vol. 29 (2022), No. 1, pp. 53–74
DOI: 10.24425/mms.2022.138548

which obtains the PSF based on the prior estimation of the motion-blurred image, and the degradation model is established to restore the image. Aiming at bypassing the shortcomings of the Hough and Radon transforms, the Gabor filter is more suitable for blur angle estimation. In addition, the blur length is estimated by the GRNN for that the nonlinear relationship between blur length and spectral characteristics. After getting blur parameters, the blurred image is restored with Wiener filter. Compared to the traditional method, it shows that the proposed method is feasible and effective in removing motion blur in tag group images.

4. Tag group recognition

4.1. Algorithm principle

Tag group recognition mainly includes two parts. One is that the tags region is roughly filtered by the Gaussian filter. Firstly, all training data set is matched by the Gaussian mixture to obtain the corresponding static background and dynamic foreground. Then, the dynamic foreground is extracted to get the rough filtered tags foreground by the Gaussian filter. The other is that tag features are extracted and screened twice by YOLOv2. The collected tag data set is pre-processed, which creates input for training and adjusting the parameters in the model repeatedly. The weights representing the tag features are generated and named in the tag foreground obtained by rough screening. Finally, the tag area is filtered again, and the specific location and range of tags are framed.

The overall flow chart of the tag group recognition algorithm YOLOv2 improved by the Gaussian filter is shown in Fig. 8.

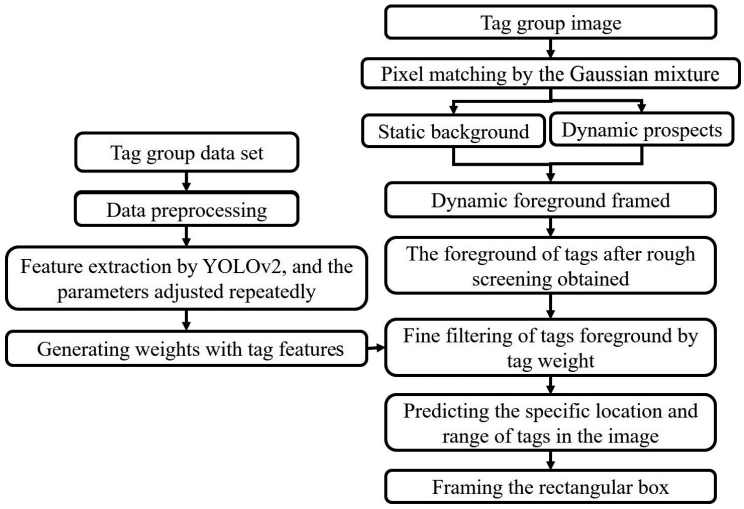


Fig. 8. Flow chart of tags group recognition algorithm YOLOv2 improved with the Gaussian filter.

4.1.1. Tags foreground extraction based on the Gaussian mixture algorithm

The adaptive Gaussian mixture model [32] is proposed by Stauffer *et al* which weights and models K Gaussian functions to represent the characteristics of each pixel in the image. After a new image is obtained, the Gaussian mixture model is updated, and each pixel in the current

image is matched with the Gaussian mixture model to separate the foreground and background. The foreground extraction method of the Gaussian mixture has a high self-adaption ability to the background, which can better describe the complex background environment. In addition, the background elimination effect of the Gaussian mixture also depends on the surrounding environment of the image and, to a certain extent, the morphological characteristics of the objects. In the actual complex environment, the relative size of the tag is small in the image, and it is easy to be affected by the surrounding environment, running speed, and direction. The average gradient of each pixel will also change accordingly, making the overall contour more complex and blur.

Therefore, based on the Gaussian mixture algorithm and the inherent characteristics of different tags, this paper improves the Gaussian mixture to adapt to the extraction of tags group foreground. The specific improvement steps are as follows:

1. The Gaussian takes the matched pixels in each image as the black foreground and the unmatched pixels as the white background. Each image is converted into a binary image.
2. For each binary image after matching, the coordinate system is established with the upper left corner of the image as the origin point and all pixels are traversed. The minimum and maximum values of x, y axis coordinates $(x_{\min}, y_{\min})$, $(x_{\max}, y_{\max})$ are found from all the black pixels, and the region of interest corresponding tags foreground is framed.
3. The *region of interest* (ROI) of the tags foreground is can be expressed as $[x_{\max} + \mu : x_{\min} - \mu y_{\max} + \beta : y_{\min} - \beta]$, which is represented by the red rectangle in Fig. 9. When $\mu = 0$ and $\beta = 0$, ROI in the tags foreground is the rectangular part of the middle solid line in Fig. 9. μ and β are adjusted by different image scenes to receive better tags area extraction effect. There are no black pixels in the matched binary image, which means that the moving foreground is not extracted, then the next image is recognized and updated.

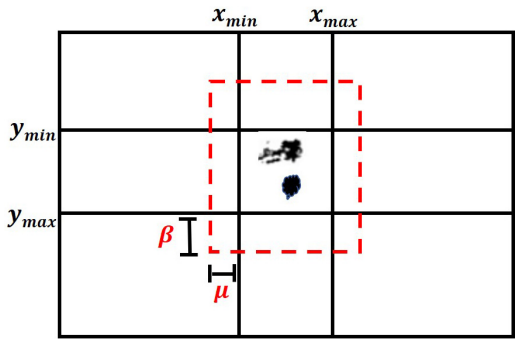


Fig. 9. Region of interest in tags foreground.

4. The coordinate information of the region of interest defined is fed back to the corresponding original image, and the tags region is extracted to complete the first coarse screening. Through the experiments of tags in different scenes, YOLOv2 improved with the Gaussian can effectively complete tag foreground extraction.

4.1.2. Tag group recognition by YOLOv2

YOLOv2 is an end-to-end object recognition algorithm which integrates region prediction and category prediction into a single neural network model [25]. It can quickly identify targets with high accuracy, which is more suitable for practical environment application. YOLOv2 has

Metrol. Meas. Syst., Vol. 29 (2022), No. 1, pp. 53–74
DOI: 10.24425/mms.2022.138548

made great improvement when compared to YOLO, *i.e.*, it adds batch normalization after each convolution layer to normalize input data of each layer [25]. It can balance the data distribution, improve the convergence speed of the model, and reduce the overfitting phenomenon.

The YOLOv2 selects the Google net as the network structure. The specific network structure is shown in Figure 10. The whole YOLOv2 model has 32 layers of neural network structure and the input image resolution reaches 416×416. The feature map of $13 \times 13 \times 1024$ is output by extracting 20 convolution layers and 5 pooling layers. Here, the convolutional kernel is 3×3 in size and the number of channels is doubled after each pooling operation. To improve the model performance, high-and-low resolution feature maps are linked by the fusion of YOLOv2 and the Gaussian. Firstly, the feature map of 16th layer $26 \times 26 \times 512$ is reconstructed into a feature map of $13 \times 13 \times 256$. Then, it is connected with the original deep feature map to form a feature map of $13 \times 13 \times 1280$. Also, different fine-grained features are used to improve the model performance, which is more conducive to the small-scale detection of tags. Then, the features are extracted by convolution and, finally, $13 \times 13 \times 13$ is the output. Tag information is classified and output on each cell of the feature map, and its specific position coordinates are predicted.

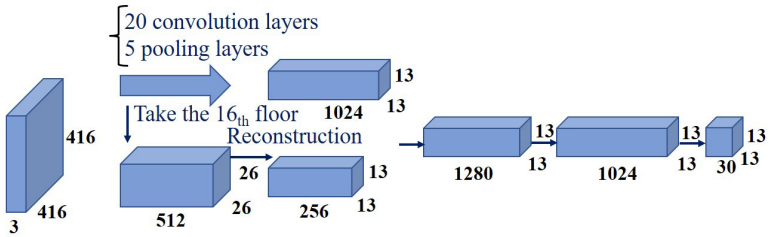


Fig. 10. YOLOv2 network structure.

To restrict the range of location prediction, YOLOv2 improves the anchor boxes to make it easier for the model to learn the location. Finally, the location information of the target is predicted for each cell of the network feature map. The specific coordinate calculation formula is shown as follows:

$$\begin{aligned} &P_r(object) * IOU(b, object) = \sigma(t_o) \\ &b_x = \sigma(t_x) + c_x \\ &b_y = \sigma(t_y) + c_y \\ &b_w = p_w e^t_w \\ &b_h = p_h e^t_h \end{aligned} \quad . \tag{8}$$

When the detected image is input, it is divided into $S \times S$ grids. If the center of the detected object is located in a grid, the grid is responsible for recognizing the object. A grid is responsible for predicting the border position of B objects which contains five normalized prediction parameters. Each bounding box predicts five values, namely t_x, t_y t_w, t_h, t_o, respectively. The first four parameters are coordinates and t_o is confidence. The distance between the cell edge and the upper left corner of the image is c_x, c_y. The length and width of the box dimension corresponding to the cell are p_w, p_h, respectively. $P_r(object)$ is the position of the anchor, and $IOU(bobject)$ is the position of the predicted object. Its purpose is to predict the location of the box by the anchor, so that the predicted box is closer to the actual object. The bounding box coordinate diagram is shown in Fig. 11. When the confidence level exceeds 25% of the threshold, the target recognition is completed by framing the candidate frame and estimating the probability of the target based on the output coordinates of the bounding boxes.

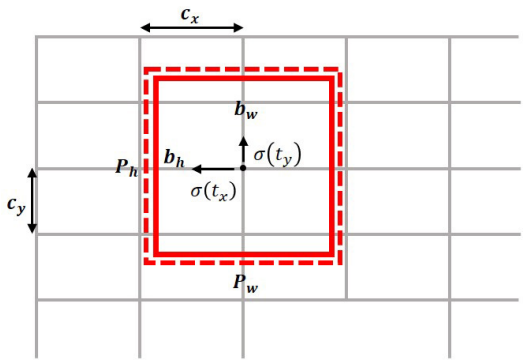


Fig. 11. Bounding box coordinate diagram.

In the RFID tag group recognition measurement system, 200 tag images of different sizes are selected in each scene as a tag group data set. Thus, in our study we have1200 tag images as the data set of the tag group in six scenes, in which 150 images are selected for each scene as the training set and the other 50 images as the test data set. The specific steps are as follows:

1. Data set pre-processing – The tag samples in the data set are filtered and denoised. Then the data set is expanded, including random clipping, rotation, changing brightness and saturation, *etc*. Tag coordinate information of each image is labelled manually. Finally, the tag image and coordinate information are used as the model input to complete the pre-processing of the data set.
2. Tag group sample training – A deep learning framework called Darknet is adopted in YOLOv2. After many times of training and adjusting parameters, the initial learning rate is set to 0.001, the size of the batch is 8, and the maximum number of iterations is 19000. The tag features are extracted and the corresponding weights are generated, which visualizes the YOLOv2 training process. When the value of the average loss function is smaller, the actual output is closer to the expected output, and the training effect is relatively better. The distribution of the average loss function is shown as Fig. 12. With the increase in the iteration, the loss function of the ordinate gradually decreases and is finally controlled within the effective range to complete the tag feature extraction and training.

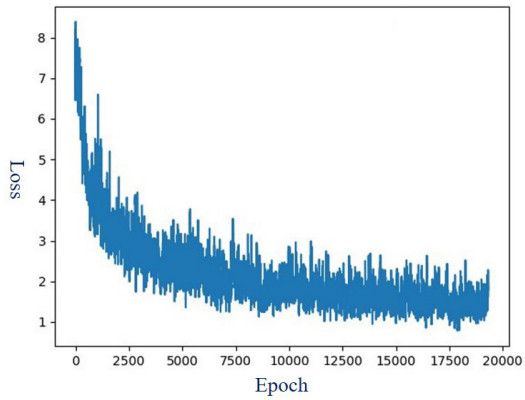


Fig. 12. Distribution of the average loss function.

Metrol. Meas. Syst., Vol. 29 (2022), No. 1, pp. 53–74
DOI: 10.24425/mms.2022.138548

In the training of different tag group data sets, YOLOv2 makes tags recognition generalization through a lot of learning and parameter optimization. It can not only verify and identify the collected data sets, but is also suitable for recognizing other untrained tag group scenes. Finally, the weight with tags feature is generated, which is screened twice in YOLOv2 improved with the Gaussian.

4.2. Analysis of experimental results

To verify the tag group recognition algorithm, this paper selects 100 images to test. The computer configuration used in the experiment is an I5-8500H, GTX1050Ti, and the running memory is 16GB. The initial learning rate is 0.001, the maximum number of iterations is 19000, the weight attenuation value is 0.005, and the batch size is set to 8. Intersection over Union (IoU) is set to 0.7. IoU represents the intersection and union ratio of the object box and the prediction box. It can not only determine positive and negative samples, but also reflects the prediction frame and detection effect. If the prediction box is close to the real box, the IoU value is larger, on the contrary, the IoU value is smaller. The calculation formula of the IoU is as follows:

$$IoU = \frac{A \cap B}{A \cup B}, \tag{9}$$

where A and B represent object data and forecast data, respectively. The IoU measures the distance and coincidence degree of the two boxes. In this paper, the rectangle box is used to recognize the specific location and range of tag area, which verifies the recognition effect of the algorithm. Fig. 13 shows the recognition effect of different algorithms on tag group images in different scenes. The first line are the tag images of different scenes. From left to right, there is one tag, two tags, three tags, five tags image, and overlapping tag group image. The second line is the effect of tag detection by YOLOv2, and the third line is the tag foreground effect image extracted by YOLOv2 improved with the Gaussian. The proposed method and YOLOv2 can effectively detect the tag position when tags are not occluded in the fifth column of Fig. 13. At the same time the confidence level of the region of interest is relatively higher than with YOLOv2. On the other hand, when the tags are occluded, YOLOv2 may misjudge the number of tags, and the YOLOv2 improved with the Gaussian can find the location of occluded tags more clearly, which can effectively complete tag coarse screening. Although the tag foreground may be mixed with

Fig. 13. Recognition effect of different algorithms with tag group images.

other moving objects and the background, it can effectively reduce the detection range of the tag as well as error detection rate of the suspicious tag area, especially in a static region.

Besides, the recognition effect of a tag group is mainly evaluated by two factors. One is accurate recognition of the tag; the other is the accurate exclusion of images without tags. In this paper, the recognition results of different algorithms are analyzed, and two evaluation criteria *Recall* and Missing Rate (MR) are selected, which are expressed as follows:

$$\begin{aligned} Recall &= \frac{TP}{AF} \\ M.R. &= \frac{FP}{FP + TN} \end{aligned}, \tag{10}$$

where TF represents the number of tags in the images identified correctly, AF is the total number of images, TN is the number of undetected tags in tag images, and FP represents the number of tag images mistakenly identified. *Recall*, *M.R.* and detection rate are used as comparison standards to verify the effectiveness of tag recognition. Recognition rate is the probability that all tags are recognized in the image. The experimental data are shown in Table 4. Here, tag-n represents the image scene with n tags randomly placed on the tray but not overlapped in the system, and its value is the average result of 50 images. Tag-com are tag group images with 5–8 tags placed on the tray and distributed in cross overlap, in which 100 images are selected for each different number of tag groups for different overlapping random arrangements. The value of tag-com is the statistical average of the 400 overlapping tag group images.

Table 4. *Recall* and *M.R.* of tags recognition on different data sets by different algorithms.

Algorithm	**Evaluation index**	**Tag-1**	**Tag-2**	**Tag-3**	**Tag-4**	**Tag-5**	**Tag-Com**
YOLOv2	Recall (%)	94.7	89.9	85.2	79.1	77.4	71.8
	M.R. (%)	5.2	1.3	1.9	2.3	2.7	3.6
	Recognition rate (%)	95.3	93.6	87.7	84.4	81.1	79.8
YOLOv3	Recall (%)	97.6	91.3	86.2	83.5	77.3	76.5
	M.R. (%)	1.5	1.6	1.1	0.9	0.7	0.8
	Recognition rate (%)	98.3	96.5	93.2	93.1	86.9	85.9
Ours	Recall (%)	98.1	92.0	86.1	81.6	76.5	74.8
	M.R. (%)	0.8	0.3	1.1	0.5	0.7	1.7
	Recognition rate (%)	97.6	95.4	93.1	93.5	88.3	87.3

To intuitively compare the advantages of this method, Table 4 compares YOLOv2 improved with the Gaussian with YOLOv2 and YOLOv3. Due to different environments and complexity in image acquisition, the recognition effects of the three methods are slightly different. With the rotation of tags on the rotating tray, the tags may overlap and block. At the same time, the tags' position is different, which may make the tags' sizes different in the image. The smaller the tags, the lower the recognition. It can be seen from Table 4 that the recognition rate of YOLOv2 improved with the Gaussian and YOLOv3 is significantly better than YOLOv2. There is little difference between the overall recognition effect of YOLOv2 improved with the Gaussian and YOLOv3. But in the complex tag group environment, our method can achieve a better recognition effect, and the recognition rate is 1.6% higher than that of YOLOv3. Thus, our method has higher robustness. In addition, the running times taken by YOLOv2, YOLOv3, and YOLOv2 improved with the Gaussian are 0.195 s, 0.101 s, and 0.113 s respectively. There is little difference in

Metrol. Meas. Syst., Vol. 29 (2022), No. 1, pp. 53–74
DOI: 10.24425/mms.2022.138548

time between YOLOv3 and Gaussian-improved YOLOv2, but YOLOv3 is a little faster. Because YOLOv2 improved with the Gaussian is more robust, our method is more in line with industrial production and practical applications in the actual complex tag group environment.

4.3. System assessment

4.3.1. Image system assessment

Compared with the previous methods [16, 31], the deep learning recognition system based on motion blur estimation and improved YOLOv2 has obvious merits. The blur and noise of polluted images are unknown and cannot be completely repaired. Since the tags are moving, it is necessary to estimate the motion blur of the tag group images and remove the sharp images. In this paper, we assume that the image known *a priori* and noise are fixed during filtering. Motion deblurring is beneficial to the recognition of subsequent tag groups. In tag group recognition, the improved YOLOv2 also shows an astonishing effect. Table 5 shows the evaluation results of tag recognition described in [16] and with the proposed method. Time refers to the average time of running 100 tag image detection. Accuracy is the number of tags images that can be correctly detected in 100 images.

Table 5. Evaluation results of tag recognition compared template matching [16] with ours.

Measure	Time (s)	Accuracy
[16]	1.416	78.6%
Ours	0.213	92.5%

We can observe the following advantages of the improved YOLOv2:

1. Template matching has its own limitations because it can only move in parallel. It is very difficult to detect the matching object that rotates or changes size in the original image. Meanwhile, boundary offsets are predicted by YOLOv2, instead of directly predicting the boundary size. When the size and rotation angle of the object change in the image, tags are recognized effectively. And this method makes image details match and recognize more clearly.
2. Template matching is able to find the object consistently with the template from an image, and the correlation coefficient of the template is obtained for each pixel. Whereas YOLOv2 receives rich information to judge the target location, size, and category. So, the accuracy of the improved YOLOv2 is 13.9% higher than that of template matching from Table 5.
3. In terms of time, the average time of template matching is 1.416s, and the computation is large and slow. Whereas, the improved YOLOv2 selects lightweight convolution, instead of standard convolution with many parameters, which improves the speed, and the time is only 0.113 s.

As a whole, the motion blur estimation and the improved YOLOv2 are embedded into the RFID tag group recognition measurement system, which can not only restore high-quality multi-task images, quickly identify all tags, but also lays a solid foundation for the performance optimization of the RFID system.

4.3.2. RFID tag group performance optimization

The purpose of RFID tag group identification is to more accurately optimize the 3D structure of the tag group and avoid communication interference between tags. Tag groups are recognized

by fusing the motion blur estimation and YOLOv2 improved with the Gaussian. Then, the coordinate values of x and y are calculated in the image acquired by the vertical camera, and the coordinate values of z are calculated in the image obtained by the horizontal camera, to find the three-dimensional coordinates of the tag group in the image. The dual CCD camera can not only restore the high-quality tag image, but also finds 3D information without massive calculations. According to [13], the relationship between the tag group in the image coordinate system and the world coordinate system is established, and then the actual position of the tag group in the world space is calculated. The system optimizes the 3D structure of the tag group more accurately through reading performance to find the optimal 3D distribution and avoid the interference of tag communication conflict. To more intuitively observe the improvement of the performance of the designed system, different methods are used to read the different numbers of tag groups with different 3D distributions. When the number of tags is different, the maximum reading distance corresponding to the tag group is compared as shown in Fig. 14.

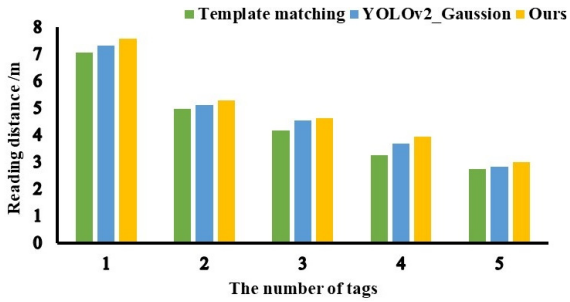


Fig. 14. Maximum reading distance to the tag group.

1–5 tags are selected in this paper to test the reading performance. The reading performance of the tag group is evaluated by the reading distance. When the tag group enters the gate in the system and all tags are read by the reader, it will trigger the laser rangefinder to read the distance from the antenna to the tag group, which is called the reading distance of the tag group. When the number of tags is the same, the previous template matching, YOLOv2 improved with Gaussian, and our method are embedded into the RFID tag group performance measurement system. The 3D distribution structure of the tag group is then changed to find the maximum reading distance. Each method runs 100 times in the system, and the average results are calculated. The ambient temperature of the semi-physical experiment simulation system is 26–28 ° and the relative humidity is 40–60%. It can be seen from Fig. 14. that the YOLO series is better than the template matching as it comes to the reading distance. Thus, it is proved that the end-to-end recognition effect is better. In addition, in the process of RFID system performance measurement, the tag movement makes the acquired tag group image blur. Compared with the only YOLOv2 improved with the Gaussian, the reading performance of our method better by at least 2.1%. Compared with the previous template matching, the reading performance is improved by at least 6.2%.

By changing the 3D structure of the tag group, the reading distance is improved. It proves that the method proposed in this paper reduces the interference between tags in the reading process, avoids the communication collision in the communication channel process, and improves the anti-collision ability of the RFID system. In future work, the optimized RFID system can be widely used in warehousing and logistics with the aim of restoring the false detection rate and improving the identification efficiency.

Metrol. Meas. Syst., Vol. 29 (2022), No. 1, pp. 53–74
DOI: 10.24425/mms.2022.138548

5. Conclusions

For more accurate identification of dynamic tag groups, this paper presents the of design a tag group dynamic recognition measurement system based on motion blur estimation and YOLOv2 improved by the Gaussian. Firstly, blur angle is estimated with the Gabor filter, and blur length is estimated by the GRNN. Secondly, the tag group region is roughly filtered by the Gaussian, tags are recognized by YOLOv2, and the tag location is screened twice. Finally, the specific location and range of tags are found. In this paper, the effective fusion of the two methods can not only restore the information of the tag group image, but also identifies the location of the tag more accurately. It helps to optimize the three-dimensional distribution of RFID tag groups, improve their corresponding reading performance and avoid the collision of tags in the process of communication. In the future, we will focus on the analysis of the factors influencing tag group recognition. For example, complex environments, different kinds of tag recognition, and abundant tags are read simultaneously. So, we will be committed to improving the detection of tag groups faster and more accurately with YOLOv4. In addition, we will collect data in the actual logistics environment and compare it with the experimental data to test the practicability and reliability of the semi-physical experimental system.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (NNSFC) (61771240) and the Six Talent Peaks Project in Jiangsu Province of China (XYDXX-058).

References

[1] Jankowski-Mihułowicz T, & Węglarski M. (2016). A Method for Measuring the Radiation Pattern of UHF RFID Transponders. *Metrology and Measurement Systems*, *23*(2), 163–172. https://doi.org/10.1515/mms-2016-0018

[2] Grover A., & Berghel H. (2011). A survey of RFID deployment and security issues. *Journal of Information Processing Systems*, *7*(4), 561–580. https://doi.org/10.3745/JIPS.2011.7.4.561

[3] Rehman, K., & Nawaz, F. (2020). Localization of Static Passive RFID Tags with Mobile Reader in Two Dimensional Tag Matrix. *Wireless Personal Communications*, *114*(1), 609–628. https://doi.org/10.1007/s11277-020-07384-1

[4] Athalye, A., Savic, V., Bolic, M., & Djuric, P. M. (2012). Novel semi-passive RFID system for indoor localization. *IEEE Sensors Journal*, *13*(2), 528–537. https://doi.org/10.1109/JSEN.2012.2220344

[5] Ni, L. M., Liu, Y., Lau, Y. C., & Patil, A. P. (2003, March). LANDMARC: Indoor location sensing using active RFID. In *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications (PerCom 2003)* (pp. 407–415). IEEE. https://doi.org/10.1109/PERCOM.2003.1192765

[6] Zhao, Y., Liu, Y., & Ni, L. M. (2007, September). VIRE: Active RFID-based localization using virtual reference elimination. In *2007 International Conference on Parallel Processing (ICPP 2007)* (pp. 56–56). IEEE. https://doi.org/10.1109/ICPP.2007.84

[7] Jin, G. Y., Lu, X. Y., & Park, M. S. (2006, June). An indoor localization mechanism using active RFID tag. In *IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC'06)* (Vol. 1, pp. 4–pp). IEEE. https://doi.org/10.1109/SUTC.2006.1636157

[8] Shen, L., Zhang, Q., Pang, J., Xu, H., Li, P., & Xue, D. (2019). ANTspin: Efficient Absolute Localization Method of RFID Tags via Spinning Antenna. *Sensors*, *19*(9), 2194. https://doi.org/10.3390/s19092194

[9] Li, M., Chen, Y., Zhang, Y., Yang, J., & Du, H. (2019, June). Fusing RFID and computer vision for occlusion-aware object identifying and tracking. In *International Conference on Wireless Algorithms, Systems, and Applications* (pp. 175–187). Springer, Cham. https://doi.org/10.1007/978-3-030-23597-0_14

[10] Duan, C., Rao, X., Yang, L., & Liu, Y. (2017, May). Fusing RFID and computer vision for fine-grained object tracking. In *IEEE INFOCOM 2017-IEEE Conference on Computer Communications* (pp. 1–9). IEEE. https://doi.org/10.1109/INFOCOM.2017.8057161

[11] Wang, Z., Xu, M., Ye, N., Wang, R., & Huang, H. (2019). RF-Focus: Computer vision-assisted region-of-interest RFID tag recognition and localization in multipath-prevalent environments. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *3*(1), 1–30. https://doi.org/10.1145/3314416

[12] Zhuang, X., Yu, X., Zhou, D., Zhao, Z., Zhang, W., Li, L., & Liu, Z. (2019). A novel 3D position measurement and structure prediction method for RFID tag group based on deep belief network. *Measurement*, *136*, 25–35. https://doi.org/10.1016/j.measurement.2018.12.071

[13] Li, L., Yu, X., Jin, Z., Zhao, Z., Zhuang, X., & Liu, Z. (2020). FDnCNN-based image denoising for multi-label localization measurement. *Measurement*, 152, 107367. https://doi.org/10.1016/j.measurement.2019.107367

[14] Li, L., Yu, X. L., Zhuang, X., Zhao, Z. M., Zhu, X. Y., Liu, Z. L., & Dong, D. B. (2020). Optimization of Radio-frequency Identification (RFID) Multi-tag Topology Based on Laser Ranging and Mind Evolutionary Algorithm (MEA). *Lasers in Engineering*, *45*(1–3), 15–34.

[15] Yu, Y., Yu, X., Zhao, Z., Qian, K., & Wang, D. (2018). Image analysis system for optimal geometric distribution of RFID tags based on flood fill and DLT. *IEEE Transactions on Instrumentation and Measurement*, *67*(4), 839–848. https://doi.org/10.1109/TIM.2017.2789122

[16] Zhuang, X., Yu, X., Zhao, Z., Wang, D., Zhang, W., Liu, Z., Lu, D. & Dong, D. (2018). A novel method for 3D measurement of RFID multi-tag network based on matching vision and wavelet. *Measurement Science and Technology*, *29*(7), 075001. https://doi.org/10.1088/1361-6501/aabcac

[17] Yu, Y., Yu, X., Zhao, Z., Liu, J., & Wang, D. (2017). Optimal Distribution of Radio Frequency Identification (RFID) Multiple Tags Based on a Support Vector Machine (SVM) and Laser Ranging. *Lasers in Engineering*, *38*(1–2), 109–124.

[18] Dobeš, M., Machala, L., & Fürst, T. (2010). Blurred image restoration: A fast method of finding the motion length and angle. *Digital Signal Processing*, *20*(6), 1677–1686. https://doi.org/10.1016/j.dsp.2010.03.012

[19] Zhuo, H. B., Bai, F. Z., & Xu, Y. X. (2020). Machine vision detection of pointer features in images of analog meter displays. *Metrology and Measurement Systems*, *27*(4), 589–599. https://doi.org/10.24425/mms.2020.134840

[20] Ta, Q. B., & Kim, J. T. (2020). Monitoring of Corroded and Loosened Bolts in Steel Structures via Deep Learning and Hough Transforms. *Sensors*, *20*(23), 6888. https://doi.org/10.3390/s20236888

[21] Shu, Y, & Gao, M. C. (2004). Restoration of the Image Blurred by Motion at Arbitrary Direction. *Computer Engineering and Applications Journal*, *31*, 361–368.

[22] Dash, R., & Majhi, B. (2014). Motion blur parameters estimation for image restoration. *Optik*, *125*(5), 1634–1640. https://doi.org/10.1016/j.ijleo.2013.09.026

Metrol. Meas. Syst., Vol. 29 (2022), No. 1, pp. 53–74
DOI: 10.24425/mms.2022.138548

[23] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 779–788). https://doi.org/10.1109/CVPR.2016.91

[24] Redmon, J., & Farhadi, A. (2017). YOLO9000: better, faster, stronger. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 7263–7271). https://doi.org/10.1109/CVPR.2017.690

[25] Kessentini, Y., Besbes, M. D., Ammar, S., & Chabbouh, A. (2019). A two-stage deep neural network for multi-norm license plate detection and recognition. *Expert Systems with Applications*, *136*, 159–170. https://doi.org/10.1016/j.eswa.2019.06.036

[26] Wang, X., Yang, L. T., Li, H., Lin, M., Han, J., & Apduhan, B. O. (2019). NQA: A nested anti-collision algorithm for RFID systems. *ACM Transactions on Embedded Computing Systems*, *18*(4), 1–21. https://doi.org/10.1145/3330139

[27] Alilou, V. K., & Yaghmaee, F. (2015). Application of GRNN neural network in non-texture image inpainting and restoration. *Pattern Recognition Letters*, *62*, 24–31. https://doi.org/10.1016/j.patrec.2015.04.020

[28] Li, A., Yang, X., Xie, Z., & Yang, C. (2019). An optimized GRNN-enabled approach for power transformer fault diagnosis. *IEEJ Transactions on Electrical and Electronic Engineering*, *14*(8), 1181–1188. https://doi.org/10.1002/tee.22916

[29] Wang, J., Chen, P., Zheng, N., Chen, B., Principe, J. C., & Wang, F. Y. (2021). Associations between MSE and SSIM as cost functions in linear decomposition with application to bit allocation for sparse coding. *Neurocomputing*, *422*, 139–149. https://doi.org/10.1016/j.neucom.2020.10.018

[30] Haouassi, S., & Wu, D. (2020). Image dehazing based on (CMTnet) cascaded multi-scale convolutional neural networks and efficient light estimation algorithm. *Applied Sciences*, *10*(3), 1190. https://doi.org/10.3390/app10031190

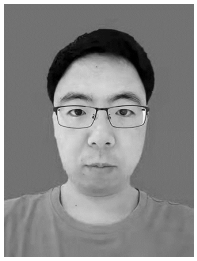
[31] Yu, X., Zhou, Y., Liu, Z., & Zhao, Z. (2019). An optimal measurement method for spatial distribution of radio frequency identification multi-tag based on image analysis and PSO. *Transactions of the Institute of Measurement and Control*, *41*(12), 3331–3339. https://doi.org/10.1177%2F0142331218823864

[32] Stauffer, C., & Grimson, W. E. L. (1999, June). Adaptive background mixture models for real-time tracking. In *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)* (Vol. 2, pp. 246–252). IEEE. https://doi.org/10.1109/CVPR.1999.784637

Lin Li received the B.Sc. in telecommunication engineering in 2012, and her M.S. Degree in control science and engineering in 2014 from Northeast Forestry University, Harbin, China. She is currently pursuing the Ph.D. degree in optical engineering at Nanjing University of Aeronautics and Astronautics. Her research interests are photoelectric measurement and instrument design.

Zhimin Zhao received the B.Sc. in 1977 and her M.Sc. in 1992 from the Nanjing University of Aeronautics and Astronautics. She was also a distinguished visiting scientist in CSIRO in Australia from 2009 to 2010. She has taken part in many research and technology transfer projects. She regularly participates in research projects as an external evaluator. She is currently Professor at the Nanjing University of Aeronautics and Astronautics. Her research topics cover a broad spectrum including computer vision, light source design, and pattern recognition.

Xiaolei Yu received the B.Sc. in 2004, and his M.Sc. in 2007 from the Nanjing University of Aeronautics and Astronautics. He received his Ph.D. degree from the University of Melbourne in 2012. He did post-doctoral research at the Nanjing University of Science and Technology, and he was rated as industrial professor of the Jiangsu Province in 2019. He currently is a postgraduate supervisor of the Nanjing University of Aeronautics and Astronautics. His research topics are mainly the Internet of Things information processing and photoelectric detection.



Zhenlu Liu received his B.Sc. degree from the Nanjing University of Aeronautics and Astronautics in 2015. He is currently pursuing the Ph.D. degree in optical engineering from the same university. His research interest include processing and RFID multitag dynamic measurement.

Ke Zhang received his B.S. degree from Jiangsu University in 2019. She is currently pursuing here M.SC. in optical engineering at the Nanjing University of Aeronautics and Astronautics. Her research interests are measurement techniques and instruments.

Shanhao Zhou received his B.Sc. degree from the China Jiliang University in 2019. He is currently pursuing the Master's degree in optical engineering at the Nanjing University of Aeronautics and Astronautics. His research interests are measurement techniques and instruments.