Multi-Stage Contour Primitive of Interest Extraction Network with Dense Direction Classification

Jinyan LU†, Quanzhen HUANG†(a), Nonmembers, and Shoubing LIU†, Student Member

SUMMARY For intelligent vision measurement, the geometric image feature extraction is an essential issue. Contour primitive of interest (CPI) means a regular-shaped contour feature lying on a target object, which is widely used for geometric calculation in vision measurement and servoing. To realize that the CPI extraction model can be flexibly applied to different novel objects, the one-shot learning based CPI extraction can be implemented with deep convolutional neural network, by using only one annotated support image to guide the CPI extraction process. In this paper, we propose a multi-stage contour primitives of interest extraction network (MS-CPieNet), which uses the multi-stage strategy to improve the discriminability of CPI and complex background. Second, the spatial non-local attention module is utilized to enhance the deep features, by globally fusing the image features with both short and long ranges. Moreover, the dense 4-direction classification is designed to obtain the normal direction of the contour, and the directions can be further used for the contour thinning post-process. The effectiveness of the proposed methods is validated by the experiments with the OCP and ROCM datasets. A 2-D measurement experiments are conducted to demonstrate the convenient application of the proposed MS-CPieNet.

key words: CPI, one-shot learning, multi-stage strategy, vision measurement

1. Introduction

Contour extraction is a widely-used computer vision task, because contour is a major feature to represent the object’s shape, structure, pose, etc. Many robotic and automation tasks are based on the image contour features [1]–[3]. Miao et al. proposed a phenotypic parameter measurement method to obtain the areas of the clustered grape grain, which has the two steps of holistically-nested edge detection and contour fitting [4]. The maximum outer contour recognition method is proposed, which uses Linear and circular contour parts to measure the pose of non-cooperative targets [5]. Cheng et al. designed a fully convolutional network for the tube contour extraction in complex scenes, which can be used for multi-exposure images [6]. The line contour segments on the objects’ end faces are extracted for the high precision representation of pose error, and the pose error can be used for automatic alignment of cylindrical objects [7]. For the monocular vision based pose estimation of manufactured metal parts, the target metal part is matched based on the line segment based geometrical representation, then EPnP is used to estimate the pose with the endpoints of the line segments [8]. Qin et al. firstly proposed the concept of contour primitive of interest (CPI), and many objects’ shapes can be represented by a set of line segment or circular arc CPIs [9]. Contour extraction is still a challenging task under complex scenes and non-standard requirements.

In recent years, more and more convolutional neural networks (CNNs) have been developed for contour extraction. Because CNN models have large receptive fields and high-dimensional hierarchical features, the contours extracted by CNN can describe the exact object shape, without involving noisy local edges. A fully convolutional encoder-decoder network is proposed to detect the general object contours [10]. Holistically-nested edge detection (HED) uses the multiple side output layers after the different stages of the VGG network, to provide the multi-scale and multi-level contour predictions [11]. Besides the related works [12]–[14], Yu et al. proposed CASEnet, which can both extract the contour and classify the contour’s category [15]. For the detection of the widely-seen line segment shaped contours, a series of line segment detection models have been designed, and the detected line segments can form the geometric representation of the image [16], [17]. Although the CNN based contour extraction has good accuracy and generalization ability, in many robotic or automation tasks, the general arbitrary contour cannot be directly used for measurement and control. Usually, one or more specific CPIs are required by a vision based task. To satisfy the non-standard vision tasks with various object categories, one/few-shot learning is very promising, to realize that a CNN model can be reused for various objects that have not been seen during training, just given one or a few annotated samples. In recent years, a series of few-shot learning based segmentation models are proposed, to realize category-agnostic semantic segmentation. A typical few-shot learning approach is based on metric learning, which uses a support image and its annotation to generate a prototype vector, then uses the prototype vector to guide the inference branch of a query image. Zhang et al. proposed the SG-One model, which uses the similarity map between the prototype and the query features to guide the segmentation branch by feature re-weighting [18]. CANet combines the prototype and the feature map, then uses convolutional layers to realize dense comparison, besides the result is refined by an iterative optimization module [19]. PANet uses cosine distance to measure the similarity between the

Manuscript received February 25, 2022.
Manuscript revised May 24, 2022.
Manuscript publicized July 6, 2022.
†The authors are with School of Electrical Information Engineering, Henan University of Engineering, Xinzheng, 451191, China.
a) E-mail: huangquanzhen666@126.com (Corresponding author)
DOI: 10.1587/transinf.2022EDP7031
prototype and the feature map, and prototype alignment regularization (PAR) was used to improve the generalization ability [20]. Feature weighting was proposed to improve the discriminativeness of the model, which scales the activations to be high in the foreground and low elsewhere [21]. Contour primitive of interest extraction network (CPieNet) is specially designed for the one-shot learning based CPI extraction task [22]. Compared to the models in [18]–[21], CPieNet is more suitable for the precise extraction of the foreground of the sharp and regular contour primitive.

Aiming to improve the CPI extraction performance, the contributions are as below.

1) We proposed a multi-stage contour primitive of interest extraction network (MS-CPieNet), considering that only one stage of comparison might not be sufficient to distinguish the entire CPI against the complex background.
2) To exploit the long-range dependencies between all the positions in the image, the spatial non-local attention module is embedded to enhance the feature extraction.
3) As an alternative to Gabor-filter based normal direction identification, the dense 4-direction classifier is designed to provide the normal directions of contour pixels, which are then used for non-maximum suppression based contour thinning.
4) MS-CPieNet outperforms CPieNet and other related models on the OCP and ROCM datasets, which demonstrates the effectiveness of the proposed methods. The convenient application of the trained MS-CPieNet, which is programming-free, is also demonstrated with a 2-D measurement experiment.

2. Methods

The one-shot learning based CPI extraction task is to obtain a CPI map $C_Q$ from an RGB query image $I_Q$, under the guidance of an RGB support image $I_S$ and its CPI annotation map $C_S$. The support and query images contain the same object, but have some differences in imaging condition. $C_Q$ and $C_S$ are binary images, in which the CPI pixels are assigned 1 and the background pixels are assigned 0. Thus, the one-shot learning based CPI extraction is described by,

$$C_Q = F(I_Q, I_S, C_S)$$

where $F$ is a CNN model for CPI extraction. As shown in Fig. 1, the line segment CPI (red) is annotated in the support image $I_S$ of a metal object. Under the guidance of the support image and annotation map, the CNN model $F$ is implemented to extract the corresponding CPI (yellow) map $C_Q$ from a query image $I_Q$.

2.1 Model Architecture

As shown in Fig. 2, the proposed MS-CPieNet is mainly composed of the feature extractor and comparison stages. The feature extractor is used to obtain a feature map from a $320 \times 320$ input image, and each pixel of the feature map corresponds to a multi-dimensional vector, that contains the high-dimensional representation of this pixel. The comparison stage uses the support feature map and the support CPI map to generate the prototype, then compares the prototype with the query feature map to output the query CPI map.

The details are given as below.

1) Feature extractor: ResNet-50 is used as the backbone [23]. The final output of ResNet-50 is processed by two $3 \times 3$ dilated convolutional layers and a $1 \times 1$ convolutional layer, in parallel. These three layers all have 128 channels and $1 \times 1$ strides. The dilated convolutional layers have the dilated rates of 2 and 4, respectively, which are used to enlarge the spatial receptive field. The outputs of the three layers are summed together and then input to a spatial non-local attention module. The low-level and mid-level features in ResNet-50 are also utilized to form the multi-scale features. The three low- and mid-level feature maps have the sizes that are $1/2$, $1/4$, and $1/8$ of the input image size, respectively. These three feature maps are processed by the three $1 \times 1$ convolutional layers, which have 16, 32, and 64 channels, respectively, and $1 \times 1$ strides. Finally, the output of the spatial non-local attention module and the three low/mid-level feature maps are concatenated together to form a 240-dimension multi-scale feature map $H$.

2) Spatial Non-Local Attention: Non-local mechanism is widely used to model the long-range pixel dependencies in videos and images [24]–[26]. For the CPI extraction task, although deep backbone and atrous convolution are utilized, the long-range dependencies existing between CPI pixels or background pixels can provide further improvement to the feature representation ability. We embed the aforementioned spatial non-local attention module to process the 128-channel multi-scale feature map $H_X$. First, $H_X$ is transformed to $H_{XA}$, $H_{XB}$, and $H_{XC}$ by three $1 \times 1$ convolutional layers, respectively, so that $H_{XA}$ and $H_{XB}$ have 64 channels, while $H_{XC}$ has 128 channels. Then, $H_{XA}$, $H_{XB}$, and $H_{XC}$, whose dimensions are $20 \times 20 \times 64$, $20 \times 20 \times 64$, and $20 \times 20 \times 128$, respectively, are reshaped to $X_A$, $X_B$, and $X_C$, whose shapes are $400 \times 64$, $400 \times 64$, and $400 \times 128$, respectively. Then the dependencies between the pixels of $X_A$ and $X_B$ are modeled as $X_D$, as given by,

$$X_{Di,j} = \sum_j \frac{\exp(X_{Ai}X_{uj}^T)}{\sum_j \exp(X_{Ai}X_{uj}^T)}$$

where $X_{Di,j}$ is the $i$th row and $j$th column element of $X_D$. $X_{Ai}$ is the $i$th row of $X_A$, $X_{uj}$ is the $j$th row of $X_B$. Then the non-local features are fused according to the short and long dependencies $X_D$, as given by,
\[ X_N = X_D X_C^T \]  

Afterwards, \( X_N \) with the shape of 400 \( \times \) 128 is reshaped back to \( H_{X_N} \) with the shape 20 \( \times \) 20 \( \times \) 128. Finally, \( H_{X_N} \) is processed by a 1 \( \times \) 1 convolutional layer and a batch normalization layer. The output of the spatial non-local attention module is the sum of \( H_{X_N} \) and \( H_X \).

3) Dense Direction Classification: We use the learnable convolutional layers to classify the normal direction, which can be adaptive to the CPIs with different scales. As shown in Fig. 2 (a), the dense direction classification is a bypass branch of the feature extractor. The two low-level feature maps are resized to 1/2 of the input image size, then concatenated together. Then three 3 \( \times \) 3 convolutional layers with the 1 \( \times \) 1 strides are used to process the low-level features. These three layers have 64, 32, and 4 channels, respectively. The softmax function is adopted for the output activation, and resized to \( D \). The values ranges among \( \{0, 1, 2, 3\} \), which indicate that the contour normal direction is closest to 0°, 45°, 90°, and 135°, respectively.

4) Comparison Stage: As shown in Fig. 2 (b) and 2 (c), the support and query feature maps \( H_S \) and \( H_Q \) are obtained from the support and query images \( I_S \) and \( I_Q \), respectively, with the shared feature extractor. Then, \( H_S \) and \( H_Q \) are processed to \( H_{SC} \) and \( H_{QC} \), respectively, by the two shared 3 \( \times \) 3 convolutional layers with the 1 \( \times \) 1 strides. These two convolutional layers both have 128 channels. Given the support annotation map \( C_S \) and the support feature map \( H_{SC} \), masked average pooling (MAP) is used to obtain a prototype \( P \), which encodes the CPI characteristics as a 128-channel vector.

\[
P_k = \frac{\sum_{i,j} H_{SC,i,j} C_{S,i,j}}{\sum_{i,j} C_{S,i,j}} \tag{4}
\]

where \( i \), \( j \), and \( k \) are the row, column, and channel indices, respectively. After obtaining the prototype, the cosine distance between \( P \) and the pixels of the query feature map is calculated by,

\[
Z_{i,j} = 1 - \left( \frac{P}{\|P\|_2} \right)^T \left( \frac{H_{QC,i,j}}{\|H_{QC,i,j}\|_2} \right) \tag{5}
\]

A sigmoid function is used to map the distance map \( Z \) to output CPI map \( C_Q \).

\[
C_{Q,i,j} = \frac{1}{1 + e^{-5(5-20Z_{i,j})}} \tag{6}
\]

Thus, when a pixel on the query feature map has a feature representation closer to the prototype, its output response is higher, which means that the pixel’s characteristics are more similar to the annotated CPI in the support image.

5) Multi-stage strategy: To improve the discriminative ability of CPI and background, we propose to compare the support image with the query image with a multi-stage strategy. The later comparison stage can exploit the former stage’s result as auxiliary knowledge, so that the CPI extraction result could be better at the later stage. As shown in Fig. 2 (c), there are \( N \) comparison stages. Each stage has its own learnable weights. As shown by the dashed arrow line in Fig. 2 (b), when the comparison stage is the 1st stage, \( C_S \) and \( C_{Q(n-1)} \) are not used to be multiplied with the feature maps. When the comparison stage is the \( n \)th \((n > 1)\) stage, the two feature maps are firstly multiplied with \( C_S \) and \( C_{Q(n-1)} \), respectively. This multiplication operation is simply realized by,
where \(i, j\) and \(k\) are the row, column and channel indices, respectively. Thus, the input feature maps of the current comparison stage can be firstly reweighted according to the former stage’s prediction. The pixels with low confidence in the former stage will be further suppressed in the current comparison stage. Thus in the current stage, the following two convolutional layers will not involve the suppressed pixels, and focuses on the further discrimination of the false positive pixels given by the former stage. The \(N\) comparison stages outputs \(N\) predicted CPI map \(C_{Q1}, C_{Q2}, \ldots, C_{QN}\). Note that the shape of \(C_{Q1}, C_{Q2}, \ldots, C_{QN}\) is 1/2 of the input image size, namely 160 \(\times\) 160. Only the final stage’s prediction \(C_{QN}\) is up-sampled by bilinear interpolation to the input size of 320 \(\times\) 320.

The query direction map \(D'_{Q}\) given by the query branch’s feature extractor and the ground truth of query direction map is labeled as \(D'_{QGT}\). Both \(D'_{Q}\) and \(D'_{QGT}\) have the shape of 320 \(\times\) 320 \(\times\) 4. The four channels indicate the confidences of the four direction classes, respectively. Only the CPI pixels are involved in the direction classification loss. Therefore, the masked cross-entropy loss is used, as given by,

\[
L_D = \frac{\sum_{i,j,c} -C_{QGT,i,j}D'_{QGT,i,j,c} \log D'_{Q,i,j,c}}{\sum_{i,j} C_{QGT,i,j}}
\]

The total loss for multi-task learning is formed by the weighted sum of the \(N\) Dice losses and the direction classification loss, as expressed by,

\[
L = \sum_{n=1}^{N} w_n L_n + \sum_{n=1}^{N} w_n + L_D
\]

where \(w_n\) is the weight of the \(n^{th}\) comparison stage.

3. Experiments and Results

3.1 Datasets

The Object Contour Primitives (OCP) dataset had 2307 images of various object categories. 1807 images with the annotations of 4844 line segments and 622 circular arcs were used for model training. 500 images with the annotations of 1297 line segments and 186 circular arcs were used for model testing. The Robotic Object Contour Measurement (ROCM) dataset was used to test the CPI extraction model in real robotic scenes, which was built with 15 3-D objects, a robot arm and an industrial camera. 523 images with the annotations of 2188 line segments and 334 circular arcs were used for evaluation. The evaluation modes included the \(I^{th}\) frame mode and the template mode. The former mode used the first frame as the support image. The remaining frames of the same scene were used as the query images, which differed with the first frame in imaging condition. Differently, the latter mode utilized a template image, which was taken with a different camera when putting the same object on a black-pad background, as the support image.

3.2 Training and Evaluation

Adam optimizer was used to train the CNN model [27]. The ResNet-50 backbone was initialized with the weights pre-trained on ImageNet. The initial learning rate, training epoch, and batch size were set to 1e-4, 40, and 4, respectively. The learning rate was halved every 10 epochs. The experiments were conducted on an NVidia RTX2080ti GPU.

The automatic support-query image pair generation was utilized to randomly generate a sample pair from one image and its CPI map, based on random image cropping, mix-up, patching, scaling, cropping, rotation, color jittering, vertical/horizontal flipping, etc. Thus the generated support image and query image had the same object and CPI, but
Comparison of Stage Number: The results were shown in Table 1. We carried out a series of ablation experiments. The weights used for loss function were selected empirically. For 2 stages, \( w_1 = 0.5, w_2 = 1.0 \). For 3 stages, \( w_1 = 0.1, w_2 = 0.5, w_3 = 1.0 \). For 4 stages, \( w_1 = 0.1, w_2 = 0.2, w_3 = 0.5, w_4 = 1.0 \). As shown by the first four rows of Table 1, the multi-stage strategy significantly outperformed the single-stage strategy, although it required the longer running time. The MS-CPieNet with 2 stages achieved the best overall MF-ODS scores. Increasing the comparison stage from 2 to 4 led to the lower performance, because the increasing model complexity might lead to the lower generalization ability.

Spatial Non-Local Attention: The 5th row shows the performance of MS-CPieNet with 2 stages and without spatial non-local attention. Comparing the 2nd and 5th rows, it was shown that the performance was significantly worse without using the spatial non-local attention module. Therefore, the spatial non-local mechanism can increase the feature extracting ability by modeling the long-range dependencies.

3.4 Comparison Experiments

The CPI extraction performance of the proposed MS-CPieNet model was compared to those of the recent related models [18]–[22], among which CPieNet is specially designed for one-shot learning based CPI extraction task. The few-shot segmentation methods [18]–[21] were re-implemented for one-shot learning based CPI extraction. All the methods were trained with the same training framework, and used MAP to obtain the prototype. Their differences mainly lied in the design of the query inference branch, and its interaction with the prototype.

As shown by the results in Table 2, the proposed MS-CPieNet with 2 stages demonstrated the best CPI extraction accuracy on both the OPC test set and the ROCM dataset. The inference time of the CNN model satisfied the real-time application’s requirement. Compared to CPieNet, MS-CPieNet mainly had three advantages. First, the multi-stage strategy could improve the model’s discriminative ability, and the false positive responses could be further distinguished in the latter comparison stage. Second, the non-local mechanism helped to capture the long-range dependences of the CPI pixels. Third, the learnable 4-direction classification was more adaptive to the different CPI scales and shapes.

3.5 Visualization Results

The CPI extraction results are visualized in Fig. 3. The MS-CPieNet model with 2 stages were used to generate these results. First, by comparing the 3rd and 4th columns in Fig. 3, it is found that the CPI maps output by the second stage had fewer false positive responses. Second, the dense direction classification branch could correctly provide the 4 discrete contour normal directions. Third, based on the 4-direction map, the single pixel wide CPI map can be obtained by contour thinning, as shown by the 5th column.

3.6 Application in 2-D Measurement

To demonstrate the application of the proposed MS-CPieNet in vision measurement, we conducted a 2-D dimension measurement experiment. The task was to inspect the external dimensions of batteries, based on an industrial camera and a telecentric lens. As shown in Fig. 4 (a), the telecentric lens was installed with the Basler acA2440-35uc industrial camera. The telecentric lens has the advantages of ignorable distortion, constant magnification and orthogonal projection. The measurement object was a dry battery. As shown in Fig. 4 (c), the telecentric lens was installed with the Basler acA2440-35uc industrial camera.
Fig. 3 Visualization of CPI extraction results with MS-CPieNet (2 stages). Each row presents the inputs and outputs corresponding to an example. In each row, the first image shows the support images, whose CPI annotation is marked by red dashed line. The second image is the query image. The third and fourth maps are the CPI maps output by the first and second comparison stages, respectively. The fifth map is the final thinned CPI map with the single-pixel width. The last map shows the dense directions, in which the 4 normal directions are marked by 4 different colors.

Fig. 4 Visual dimension measurement experiment setup. (a) 2-D measurement system. (b) Measurement object. (c) Measurement process.

measure $\Phi_E$, a support image $I_{SE}$ of the battery's end-face was saved, and the inner circle of the plastic coat was manually labeled to obtain the support CPI map $C_{SE}$. The camera captured a query image $I_{QE}$ of the battery's end-face. The proposed MS-CPieNet was used to extract the inner circle from $I_{QE}$ under the guidance of $I_{SE}$, and output the CPI map $C_{QE}$. Then the ellipse fitting was used to obtain the accurate parametric equation of the inner circle, and the long axis length of the fitted ellipse was regarded as the diameter $\varphi_E$ of the required inner circle in the image space. Finally, due to the constant magnification of telecentric lens, we simply multiply $\varphi_E$ with the pixel equivalent to obtain the measured diameter $\Phi_E$ in the Cartesian space. Similarly, to measure $\Phi_S$, a support image $I_{SS}$ of the battery body's side view was saved, and the two side lines of the battery body were manually labeled to obtain the support CPI map $C_{SS}$. The camera captured a query image $I_{QS}$ of the battery body's side view. The proposed MS-CPieNet was used to extract the two side line segment contours from $I_{QS}$ under the guidance of $I_{SS}$, and output the CPI map $C_{QS}$. Then the line fitting was used to obtain the accurate parametric equation of the two side lines, based on which the distance $\varphi_S$ of the parallel side lines was calculated in the image space. Finally, we multiply $\varphi_S$ with the pixel equivalent to obtain the measured diameter $\Phi_S$ in the Cartesian space.

The pixel equivalent was calibrated as $0.0067\text{mm/pixel}$. The raw image resolution was $2048 \times 2048$, and the image was resized to $320 \times 320$ before inputting to MS-CPieNet. The measurement was repeated for 5 times. The battery position, battery orientation, and the lighting condition were randomly changed in a limited range during the measurement process. Using a digital caliper, the ground truths of $\Phi_E$ and $\Phi_S$ were obtained as 6.60mm and 10.31mm, respectively. As a result, the dimension measurement result was shown in Table 3. The maximum dimension measurement error was within 0.02mm, which demonstrates a good accuracy for inspecting the battery’s outer dimensions. A limitation of the measurement accuracy was caused by the limited resolution of the MS-CPieNet’s input. The CPI extraction results with MS-CPieNet and the measured battery were visualized in Fig. 5. The application of the proposed MS-CPieNet was convenient and flexible, because for the specific measure-

Table 3 Dimension measurement result.

<table>
<thead>
<tr>
<th>No.</th>
<th>Measured $\Phi_E$ (mm)</th>
<th>Error $E_E$ (mm)</th>
<th>Measured $\Phi_S$ (mm)</th>
<th>Error $E_S$ (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.609</td>
<td>0.009</td>
<td>10.314</td>
<td>0.004</td>
</tr>
<tr>
<td>2</td>
<td>6.607</td>
<td>0.007</td>
<td>10.326</td>
<td>0.016</td>
</tr>
<tr>
<td>3</td>
<td>6.615</td>
<td>0.015</td>
<td>10.301</td>
<td>-0.009</td>
</tr>
<tr>
<td>4</td>
<td>6.592</td>
<td>-0.008</td>
<td>10.321</td>
<td>0.011</td>
</tr>
<tr>
<td>5</td>
<td>6.602</td>
<td>0.002</td>
<td>10.298</td>
<td>-0.012</td>
</tr>
</tbody>
</table>
ment task in this experiment, we only needed to manually set the support images and CPI annotations. No extra programming and tuning were required in the image feature extraction step.

4. Discussion and Conclusion

In this paper, to realize the better one-shot learning based CPI extraction, we propose a multi-stage contour primitives of interest extraction network (MS-CPieNet). First, the multi-stage comparison strategy is designed to improve the model’s discrimination ability of CPI pixels. Compared to the traditional single stage comparison, the multi-stage strategy can effectively reduce the false positive outputs. Second, the spatial non-local attention mechanism is embedded in the feature extractor, so that the long-range dependencies of the CPI pixels can be captured and a global spatial attention is realized. Third, for the contour thinning post-process, the dense direction classification is utilized to provide the 4 discrete normal directions of the CPI pixels, which is more adaptive to the different CPI shapes and sizes. The learning of MS-CPieNet is supervised by the combination of the multiple Dice losses and a masked cross-entropy loss. The proposed MS-CPieNet outperforms CPieNet on both OCP and ROCM datasets in CPI extraction accuracy, also presents the flexibility and convenience in practical application.

However, our model has some limitations because one-shot learning is based on the similarity comparison between query and support images. First, the model has limited robustness to orientation changes. Large orientation change will cause very low similarity and the similarity-based guidance will not work. Then, the model cannot distinguish the target object with a very similar object, because their appearance features are very close. Fortunately, in industrial structured scenes, the large uncertainty and cluttering scene can be avoided by controlling the environment and pre-processing the image. In the future, it is promising to realize few-shot learning to overcome the inherit limitation of one-shot learning, improve the robustness of our method and expand its application scenario.

Acknowledgments

This work was supported by the National Natural Science Foundation (62173126); Program for University Scientific Innovation Team of Henan Province (21IRTSTHN017); Key R & D and Promotion Projects of Henan Province (202102210187, 222102210268); Key scientific research projects of colleges and universities of Henan Province (22A470001); National College Student Innovation and Entrepreneurship Training Program (202111517017).

References


Jinyan Lu received the PhD degree from the Institute of Automation, Chinese Academy of Sciences (IACAS), Beijing, in 2016. She is currently a lecturer in School of Electrical Information Engineering, Henan University of Engineering. Her research interests include visual measurement, robot control, and intelligent control.

Quanzhen Huang received the PhD degree in Control Science from Shanghai University, China, in 2012. He is a Professor in School of Electrical Information Engineering, Henan University of Engineering. His research interests include structure vibration control and advanced intelligent control.

Shoubing Liu received the PhD degree from Donghua University, Shanghai, in 2018. He is currently a lecturer in School of Electrical Information Engineering, Henan University of Engineering. His research interests include SAW device, sensor and instruments.