

The Optimized Intelligent Algorithms on Face Recognition and Tracking for ROS-based Robots

Yue Chen, Shuhao Tian, Huailin Zhao, Shengyang Lu

School of Electrical and Electronic Engineering, Shanghai Institute of Technology, No. 100, Haiquan Road, Shanghai 201418, China

E-mail: 405838753@qq.com

<https://www.sit.edu.cn/>

Abstract

With the development of artificial intelligence, face recognition and tracking technology have been widely used in many fields such as target positioning, automatic driving, and human-computer interaction. Recently, a large number of face detection, recognition and tracking algorithms have emerged, but there are still many shortcomings in practical applications, such as slow face detection, low detection accuracy, and face recognition and tracking for ROS robots Algorithms are rare. This paper improves the traditional Haar-like algorithm and LK optical flow tracking algorithm, and designs a ROS robot platform based on the improved algorithm. By comparing the accuracy and timeliness of face detection and tracking between the improved algorithm and the traditional algorithm, the superiority of this design algorithm is obtained.

Keyword: Face Tracking and recognition; ROS robots; LK optical flow; Haar-like algorithm

1.Introduction

Video tracking[1] is an important branch of computer vision. With the increasing impact of artificial intelligence on human production and life, video tracking is widely used in many fields such as target positioning, video tracking, automatic driving, and human-computer interaction. Recognition and tracking of mobile faces is an important technology to improve the intelligent level of mobile robots. Researchers

have developed a variety of computer face recognition and tracking algorithms. The ultimate goal is to solve noise interference, target occlusion of rapid movement and the surrounding environment More complex issues and real-time tracking. With the continuous development of computer technology, a large number of algorithms for real-time video tracking have emerged, mainly including three types of algorithms based on target external characteristics, target contours, and target local

areas. Face detection and tracking is an extremely important research field in video tracking technology, and it is the basis of many face applications. Therefore, research, processing and tracking of faces is a very meaningful direction in face video surveillance. However, in practical applications, face detection technology still has some shortcomings, such as slow face detection speed, low accuracy of face detection, and face recognition and tracking systems for ROS robots are rare. This paper designs a set of face recognition and tracking system suitable for ROS mobile robots to improve the intelligence of ROS robots.

Ebied [2] proposed the application of Principal Component Analysis (PCA) in face feature extraction and dimensionality reduction to solve the problem of reduced face recognition rate caused by the large dimension of face features. This method applies mirroring technology to the parity analysis of feature maps. The final simulation results show that this method effectively improves the accuracy of face detection and face recognition. But the algorithm runs slower and has higher requirements on the training set. TFCootes[3] and his research team proposed an Active Shape Model (ASM) based on statistical parameterization. This method builds a statistical parameterized model based on the statistical analysis results of the target contour feature vector, which can reflect Draw out the law of target shape change. Because of its simplicity and high efficiency, the algorithm has been well applied in face detection and tracking. Dejun Tang from the School of Computer Science and Technology of Dalian Maritime University^[4] studied the image feature extraction and matching technology in face recognition, using the improved Singular Value Decomposition (SVD) method of class estimation space for feature extraction to solve face recognition. In the process, the singular value vector is directly used as the face feature, which leads to the problem of low recognition rate. In addition, in order to reduce the instability of face recognition caused by changes in external factors such as illumination, expression, noise, posture and other factors, a face feature representation method that combines multi-scale global features and local features is proposed, and then rough integration is used. The simplified algorithm performs feature selection on the extracted features. Through the above

overview and analysis of existing face detection and face recognition algorithms, the existing methods have achieved certain success in face recognition and detection in complex environments and occlusions, but there is still much room for improvement. It is very important to further improve the accuracy and robustness of face recognition and detection under complex conditions.

In the process of realizing mobile face recognition and tracking, there are mainly three issues involved: 1) Recognizing the face, 2) Marking the key points of the face, 3) Detecting the key points before and after the face moves, and judging whether it is the same face. Achieve correct tracking. Therefore, after analyzing and comparing most face recognition algorithms, this design improves the Haar-like[5] face recognition algorithm to improve the accuracy and speed of face recognition. Then use the improved LK optical flow tracking algorithm[6] to track a specific moving face by detecting the key points before and after the face moves. Finally, the Camshift[7] color tracking algorithm is used for face tracking, and the improved algorithm is compared to realize the face recognition and tracking of the ROS robot. The main innovations of this article are as follows:

- (i) When marking the key points of the face, we have improved on the basis of the Harris algorithm for corner detection and adopted a more accurate marking method, which is called the better key point marking algorithm.
- (ii) This paper uses an improved LK optical flow tracking algorithm to detect the key points before and after the face moves, so as to realize the tracking of a specific moving face. Reduce the impact of the surrounding environment on face tracking.
- (iii) As a control, we also applied the Camshift color tracking algorithm to research and experiment, and compared and analyzed the experimental results of the LK optical flow tracking algorithm.
- (iv) This paper designs and implements the face recognition and tracking system of the ROS mobile robot platform, on the ROS robot platform, which can meet the accuracy and real-time performance of face recognition and tracking in complex environments through real environment tests.

2. Face recognition and tracking algorithm design

Taking into account the recognition speed, accuracy and compatibility with the robot system, the face recognition and tracking design system is mainly composed of Haar-like feature recognition, better key point marking and adding, and LK optical flow tracking. We call it the LK optical flow tracking scheme.

2.1 Face recognition based on Haar-like detector

Haar-like classifier detection target is to use the features of known training samples to compare features with unknown images to achieve target detection. This article is mainly used in face detection, generally using two rectangular features, by making the difference between the white and black area pixels to meet the detection needs. The Haar-like detector is composed of Haar-like features, Haar integral map, and cascade classifier.

Haar-like features can be roughly divided into three types: edge features, linear features, central features and diagonal features. According to the above-mentioned feature group, the Haar-like detector consists of black and white rectangles as feature templates, and its feature value is defined as the difference between the white rectangular pixels and the black rectangular pixels. The Haar-like detector uses the feature template to arbitrarily change its specifications in the image, process the image, and use the exhaustive method to calculate the feature value of the face area and the feature value of the non-face area, and judge the image gray by the feature value Degree changes and detect target recognition. The application effect is shown in Fig.1:

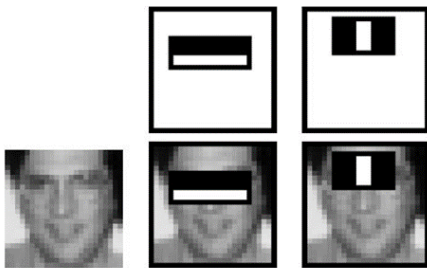


Fig.1 Haar module face recognition application

The Haar-like detector uses a sliding window to

detect the area of the face. In order to calculate the pixel difference between the feature maps, OpenCV is used to grayscale the RGB feature maps detected by the Haar detector and normalize them. The pixel value of each sliding window is grayed out. For each pixel $i(x, y)$, the average gray value of the pixel value $mean$ and the square average gray value sq_{mean} are calculated as follows:

$$mean = \frac{\sum_{x=0}^w \sum_{y=0}^h i(x, y)}{w \bullet h} \quad (1)$$

$$sq_{mean} = \frac{\sum_{x=0}^w \sum_{y=0}^h i^2(x, y)}{w \bullet h} \quad (2)$$

Then the normalization factor (varNormFactor) and the normalized haar feature value (normValue) are obtained by normalizing the image. The calculation formula is as follows:

$$var NormFactor = \sqrt{sq_{mean} - mean^2} \quad (3)$$

$$normValue = \frac{featureValue}{var NormFactor} \quad (4)$$

Where $featureValue$ is the feature value of each pixel in the feature map. In order to improve the effect of face detection, this paper adopts a cascade classifier called *Adaboost*, which is achieved by cascading multiple weak classifiers to form a strong classifier. The specific algorithm of the *Adaboost* cascade classifier is as follows:

- The picture to be input is preprocessed to obtain the characteristic value corresponding to each pixel of the picture.
- Use the Haar classifier to calculate the Haar feature value ($normValue$) corresponding to the feature value and compare it with the No. 1 face feature. If not, go to step(f), if yes, go to step(c).
- Compare the Haar feature value ($normValue$) with the No. 2 face feature, if it does not match, go to step(f), if it matches, go to step(d).
- Compare the Haar feature value ($normValue$) with the No. 3 face feature, if it does not match, go to step(f), if it matches, go to step(e).
- Confirm that it is a human face.

(f) Confirm that it is not a human face.
The specific algorithm flow chart is shown in Fig. 2:

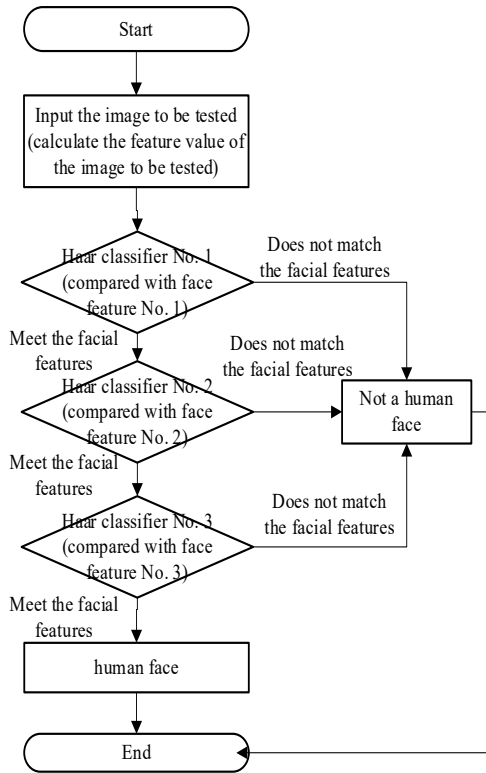


Fig.2 Flow chart of cascade classifier judgment

According to the results of face detection at each level, the image area similar to the face is screened to accurately locate the face. The scalefactor variable parameter is introduced to set the window size. The smaller the value of this parameter, the smaller the size of the sliding window and the more detailed the scan of the face, but it will take more time and memory to scan.

2.2 Better key point marking algorithm

The better key point marking algorithm mainly uses the gray-scale change relationship between the pixels between the marked corner points. The corner points are mainly the connecting points of the contour lines of the objects in the image, as shown in Fig.3. Use a fixed window to slide in any direction on the image to be tested, and compare the degree of pixel grayscale changes in the window before and after the slide. If there is a large gray scale change, then it is determined that the window has corners.

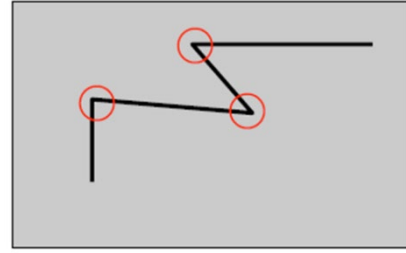


Fig.3 Corner detection

Use corner detection[9] scoring function for better key point marking algorithm, the specific scoring function is as follows

$$R = \min(\lambda_1, \lambda_2) \quad (5)$$

Manually set two feature values λ_1 and λ_2 . When the score of the picture to be detected exceeds the threshold values λ_1 min and λ_2 min, it will be judged as a corner point. The specific application of ST scoring function is shown in Fig.4.

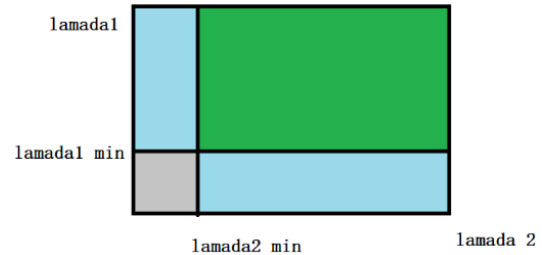


Fig.4 ST scoring function to judge corner points

After multiple tests, the appropriate number of corner points is selected, and the selection of key points is determined by the number of corner points, so that the face can be more accurately recognized and detected.

2.3 LK optical flow tracking algorithm

After extracting the key point features, the LK optical flow tracking algorithm [11] is used for real-time face tracking.

We assume:

- (1) The brightness is constant;
- (2) The time is continuous or the motion is "differential motion" and the space is consistent;
- (3) The adjacent pixels remain adjacent and have similar motions.

According to the above three assumptions, there is a picture at time t , the internal coordinates of the picture are a point at the position (x, y) , and the pixel gray level is $I(x, y, t)$; according to the assumption (2) after differential motion, That is, after moving (dx, dy) distance after dt , its gray value is $I(x+dx, y+dy, t+dt)$; finally, according

to assumption (1), the previous t time and $t+dt$. The gray scale is the same at all times. The mathematical expression is as follows:

$$I(x, y, t) = I(x+dx, y+dy, t+dt) \quad (6)$$

After differential transformation and equation transformation, the following formula is obtained:

$$\begin{bmatrix} \frac{dI_1}{dx}, \frac{dI_1}{dy} \\ \dots\dots\dots \\ \frac{dI_k}{dx}, \frac{dI_k}{dy} \end{bmatrix} \begin{bmatrix} \frac{dx}{dt} \\ \frac{dy}{dt} \end{bmatrix} = \begin{bmatrix} \frac{dI_1}{dt} \\ \dots\dots\dots \\ \frac{dI_k}{dt} \end{bmatrix} \quad (7)$$

Among them, dI/dt represents the gray changes around t and $t+dt$ at the two moments, dx/dt and dy/dt , represents the movement speed of the point in the axis x , y direction, and dI/dx and dI/dy represents the gray changes in the axis x , y direction.

Although the equation cannot be solved based on the position information of a single key point, the position of the key point in the picture after the elapse of time dt can be obtained by combining multiple key point information and hypothesis (3).

In the algorithm design, this design takes LK optical flow tracking as an extension of the better key point. Therefore, the LK optical flow tracking node can be used to subscribe to the key points that have been obtained in the better key point node, and the algorithm can be used to pass the current key point. And two grayscale pictures to predict the next set of key points, and perform backward prediction.

2.4. Optimal design of LK optical flow tracking algorithm

In the process of key point tracking, since the key points of the tracking are judged by the ST scoring function to have low scores, these key points are automatically discarded by the optical flow tracker. This phenomenon leads to the loss of faces during the tracking process. And misjudgment. Therefore, the algorithm is improved and optimized for the above problems. In the process of tracking, use the tracked results to make a statistical set, search for each frame in the video tracking process, and count, delete the abnormal feature points, update the statistical set, and search for the next frame. The specific flow

chart is as follows:

- Make a data set based on the data points obtained by searching for the first k frames;
- Search for $k+1$ frames, compare and distinguish the feature points. If it is determined to be an abnormal point, go to step(c). If it is not an abnormal point, go to step(d);
- Remove anomalies;
- Keep feature points;
- Judge whether it is the last frame, if it is the end of the last frame, if it is not the last frame, go to step(f);
- Update the statistics set.

Based on the above algorithm, the detection accuracy of the LK optical flow algorithm can be improved, and the impact of this method on the complex environment can also be minimized. In order to improve the accuracy of the recognition and tracking algorithm, this algorithm is used to improve the LK optical flow algorithm. Although this algorithm has a certain increase in the amount of calculation, the accuracy and robustness have been improved a lot. The specific algorithm flow chart is shown in Fig. 5 below:

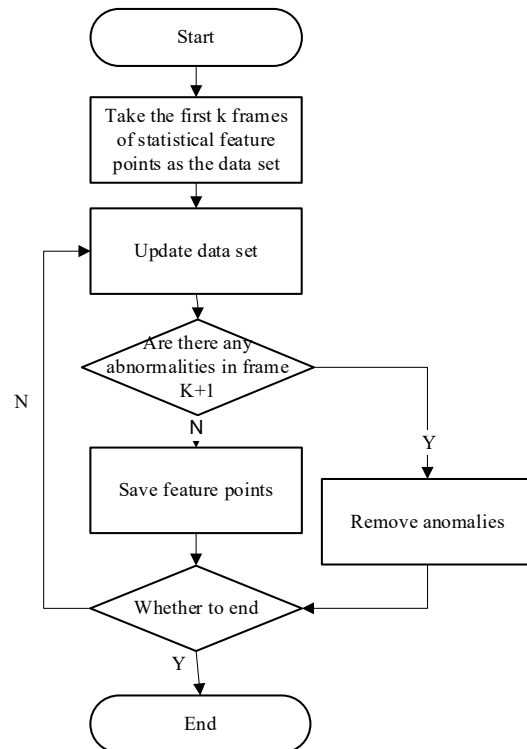


Fig.5 Optical flow statistics algorithm flow chart

3. Design of ROS robot platform based on face recognition and tracking algorithm

The experimental environment is Ubuntu system, which is realized by programming under the terminal. The hardware platform is Turtlebot2 robot, Kinect camera and host computer. After programming the upper computer loaded with the Ubuntu system and completing the basic face recognition and tracking algorithm design, the corresponding Python program was designed and tested on the ROS robot Turtlebot2.

3.1 Feature point tracking test

Process the video information received by Kinect, recognize the face and track the face in real time, and finally achieve the purpose of face recognition and tracking in a complex environment. Fig. 6 is the hardware platform built for this experiment.

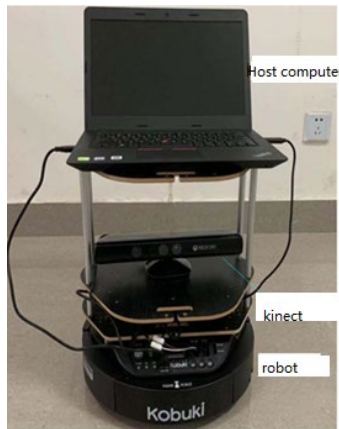


Fig.6 The mobile robot Turtlebot2

The system uses the traditional Harr-like algorithm and LK optical flow algorithm to meet the basic requirements of face recognition and tracking, but there are some problems. As shown in Fig. 7, when the face recognition and tracking program runs for a certain period of time, some of the key points will drift from the ROI area (the face recognition area) to other non-target areas. And as the running time goes by, the number of key points in the recognition screen is significantly reduced, until a few key points are left in the face area, which cannot allow the robot to track well.



Fig.7(a) 103th frame Fig.7(b) 107th frame

Fig.7 Face tracking effect before optimization

In response to the above problems, we modified and optimized the algorithm. Fig.8 uses the algorithm proposed in this article. It can be seen from Fig.8 that with the increase of time, the size and shape of the frame tracked by the face in the process of continuous movement remain unchanged, and the tracked feature points are also relatively increased and successfully completed the tracking. It can be seen that compared with the traditional algorithms, the algorithm proposed in this paper greatly reduces background information, makes full use of target information, and effectively reduces false feature points. Not only that, it also ensures that the number of feature points makes the tracking accuracy and robustness greatly improved.

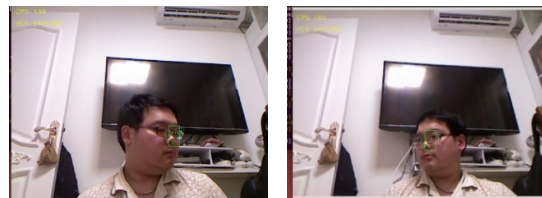


Fig.8(a) 139th frame Fig.8(b) 160th frame

Fig.8 Face tracking effect after optimization

As shown in Fig.8 above, the improved face recognition and tracking algorithm in the ROS robot platform has a greatly improved tracking effect compared with the previous one.

3.2 Face recognition test

The ROS robot platform designed this time can achieve the tracking test of the human face, that is, when the human face is moving, the robot also follows the human face to rotate and detect and recognize the human face. In order to verify the accuracy of the face recognition of the robot in the process of moving, the vertical upward is adopted, and the left and right swings at the same speed are used to measure the maximum detectable angle 0° of the robot. As shown in Fig.

9 below:



Fig.9(a) 0°

Fig.9(b) 15°

Fig.9 Swing tracking process

From the above figure, it can be seen that when the human face rotates at a certain angle, the

robot is also rotating with the human face. In order to obtain the tracking range of the robot, the human rotation angle is used as a reference for comparative analysis. In this experiment, the traditional Haar-like+optical flow tracking algorithm is compared with the improved Haar-like+optical flow tracking algorithm and the Camshift algorithm, and the superiority of this algorithm is obtained. The comparison is shown in Table 1:

Table 1 Algorithm effect comparison table

Algorithm name	Face swing angle	Recognition speed(Frames/sec)
Traditional Haar-like + optical flow tracking	$\pm 10^\circ \square \pm 15^\circ$	92
Camshift tracking algorithm	$\pm 24^\circ \square \pm 26^\circ$	72
Ours	$\pm 26^\circ \square \pm 28^\circ$	87

From the above table, it can be found that the detection angle of the algorithm in this paper has been greatly improved compared with the traditional Haar-like+optical flow tracking algorithm, but the detection speed has been reduced due to the increase in calculation. But compared to the Camshift tracking algorithm, both the accuracy and the detection speed are improved. The algorithm in this paper has a good effect.

4. Conclusion

This paper improves the traditional Haar-like algorithm and optical flow tracking, and designs and optimizes a set of face recognition and tracking system based on ROS mobile robot. The superiority of the improved algorithm is obtained by comparing with the traditional algorithm. The specific work is summarized as follows:

1. Implement and improve the traditional Haar-like algorithm and LK optical flow tracking algorithm.
2. Designed a face tracking platform based on the ROS system, and compared the improved algorithm with the traditional algorithm, and got the following conclusions:
 - A. The processing accuracy of the traditional Haar-like algorithm+optical flow tracking algorithm is very different from the algorithm in this paper, but the speed is faster, and the Camshift algorithm is not as good as the

algorithm in both processing accuracy and speed.

B. The algorithm in this paper is applied to the ROS robot platform, and the accuracy of robot face recognition is greatly improved compared with traditional algorithms.

5. References

1. Schnell Anna Elisabeth, Van den Bergh Gert, Vermaercke Ben, Gijbels Kim, Bossens Christophe, de Beeck Hans Op. Face categorization and behavioral templates in rats[J]. Journal of vision, 2019, 19(14).
2. Chen Jeng. A new patch-based LBP with adaptive weights for gender classification of human face[J]. Journal of the Chinese Institute of Engineers, 2020, 43(5).
3. Zhang G, Vela P A . Good features to track for visual SLAM[C]// IEEE Conference on Computer Vision & Pattern Recognition. IEEE Computer Society, 2015.
4. Xiaochen L, Xiaojie L, Yufeng X, et al. Evaluation of Lucas-Kanade based optical flow algorithm[C]// 2018 IEEE CSAA Guidance, Navigation and Control Conference (GNCC). IEEE, 2018.
5. Li Wang, Huailin Zhao, Yaoyao Li. Research on Research on the multi-scale network crowd density estimation algorithm based on attention mechanism. The 2019 International Conference on Intelligent Informatics and BioMedical Science (ICIIBMS2019), Nov.21-24, 2019, Shanghai, China, pp272-278.
6. Yawei Hou, Huailin Zhao. Handwritten Digit Recognition Based on Depth Neural Network. 2017 International Conference on Intelligent

- Informatics and BioMedical Science (ICIIBMS2017), Nov.24-26, 2017, Okinawa, Japan. pp35-38. (EI)
7. Shunzhou Wang, Huailin Zhao, Weiren Wang, Huijun Di, Xueming Shu. Improving Deep Crowd Density Estimation via Pre-Classification of Density. The 24th International Conference on Neural Information Processing (ICONIP2017), Nov.14-18, 2017, Guangzhou, China. pp260-269. (EI)
8. 8.Muhammad Asif, Nadeem Haider, Qasem Al-Mdallal, et al. A Haar wavelet collocation approach for solving one and two-dimensional second-order linear and nonlinear hyperbolic telegraph equations. 2020, 36(6):1962-1981
9. Multimedia; New Multimedia Study Findings Reported from National Taipei University (Fast Iris Localization Using Haar-like Features and Adaboost Algorithm). 2020,
10. Xirui Ma, Yizhou Lin, Zhenhua Nie et al. Structural damage identification based on unsupervised feature-extraction via Variational Auto-encoder[J] Measurement, 2020, 160
11. Guangqiang Li, Lei Yu, Shumin Fei. A Binocular MSCKF-Based Visual Inertial Odometry System Using LK Optical Flow. 2020, :1-16.
12. Multimedia; Investigators at Jiangsu University Describe Findings in Multimedia (Fast Stereo Visual Odometry Based On Lk Optical Flow and Orb-slam2). 2020.
13. Timotej Gašpar, Miha Deniša, Primož Radanovič, et al. Smart hardware integration with advanced robot programming technologies for efficient reconfiguration of robot workcells. 2020, 66
14. 고동현, 이재현, 김정인, et al. Integration of Mobile Robot and Manipulator Using the ROS. 2020, 29(3):182-188.
15. Iván del Pino, Miguel Á. Muñoz-Bañón, Saúl Cova-Rocamora, et al. Deeper in BLUE Development of a roBot for Localization in Unstructured Environments. 2020, 98(6):207-225.
16. Enrique Coronado, Gentiane Venture. Towards IoT-Aided Human-Robot Interaction Using NEP and ROS: A Platform-Independent, Accessible and Distributed Approach. 2020, 20(5)
17. Umberto Maniscalco, Antonio Messina, Pietro Storniolo. ASS4HR — An Artificial Somatosensory System for a Humanoid Robot. The ROS package. 2020, 11
18. Optical Research; New Data from China Nuclear Power Engineering Co. Ltd. Illuminate Findings in Optical Research (High Precision Indoor Positioning Method Based On Visible Light Communication Using Improved Camshift Tracking Algorithm). 2020, :337-.
19. Optical Research; New Data from China Nuclear Power Engineering Co. Ltd. Illuminate Findings in Optical Research (High Precision Indoor Positioning Method Based On Visible Light Communication Using Improved Camshift Tracking Algorithm). 2020,
20. Wanchao Mao, Hongyun Xie, Zequn Tan, et al. High precision indoor positioning method based on visible light communication using improved Camshift tracking algorithm. 2020, 468
21. Zhang Xiangnan, Gong Wenwen, He Qifeng, et al. Camshift tracking method based on correlation probability graph for model pig. 2020, 2020(1)
22. Gao Chaoqian, Chen Hu, Li Tianping. Target Tracking Based on Camshift Algorithm and Multi. 2020, 1550(3)
23. 양병호, 이명진. Study on the Video Stabilizer based on a Triplet CNN and Training Dataset Synthesis. 2020, 25(3):428-438.Ruiheng Zhang, Lingxiang Wu, Yukun Yang, et al.
24. Multi-camera multi-player tracking with deep player identification in sports video. 2020, 1020tabek Khudayberdiev, Muhammad Hassaan Farooq Butt.
25. Fire detection in Surveillance Videos using a combination with PCA and CNN. 2020, 3(3).