

Fast Template Matching Algorithm Based on Contour Hu Invariant Moments Object Recognition Method

Lingmin Yang, Kongyang Jiang*, Chaoyong Fang, Long Chen, Chunying Zhang, Zhongyu Wang, Menghui Han, Liuyu Hao

School of Mechanical Engineering, Liuzhou Institute of Technology, Liuzhou, Guangxi, 545026, China

**Corresponding author*

Abstract: *The object recognition of the vision robot was studied in order to address the problem of how an industrial robot can rapidly and autonomously identify and grasp a specified target workpiece under multi-target conditions, and to apply the vision robot object recognition technology to the design of an intelligent grasping system for industrial robots. The research method is based on the pattern recognition of the image, the detection and positioning research, the relationship between the visual image and the workpiece positioning and grasping, and the proposed monocular vision grasping system based on the contour Hu-invariant moment fast template matching algorithm. The image acquired by the camera is pre-processed first, then the target workpiece is identified by the contour Hu-invariant moment template matching algorithm, the position of the identified target workpiece is obtained by using the principle of minimization of contour moment and second order moment of inertia, and finally the obtained position is sent to the robotic arm control system to guide the robotic arm grasping by establishing SO CKET communication. Based on DobotStudio, DobotVisionStudio 1.4.2, MVS software development platform and DobotStudio robot, the intelligent gripping system was built and tested.*

Keywords: *Machine recognition; item recognition; fruit classification*

1. Introduction

Robot vision has always been a popular topic of research in the field of robotics. With the improvement of technology, more and more robots will be used in security systems, industrial processing, aerospace and other fields, and many application scenarios require robot vision systems with the ability to identify, locate and track targets. This project summarizes the experience and results of previous work, and mainly does the following work: First, the principle of binocular vision system is studied. The linear imaging model of the camera under ideal conditions is analyzed, and the calibration of the camera parameters considering camera aberrations is also introduced. In addition, Matlab simulations based on the existing algorithms are carried out to reconstruct the filmed scenes and to determine the accuracy of the camera parameters by inverting the positions of the target feature points according to the experimental parameters. Second, the target matching algorithm of the vision system was studied. The application of SIFT (scale-invariant feature transform) in target matching is studied, and the original RANSAC algorithm is improved to eliminate the wrong matching points. The BRIEF (Binary Robust Independent Element Feature) algorithm is studied; the above mentioned matching algorithms are compared and analyzed, and it is concluded that the BRIEF algorithm has the best real-time performance, but cannot handle the image rotation problem. Again, the current mainstream tracking algorithms are studied. The color probability density, back-projection and Mean Shift modules of the Camshift algorithm are analyzed, and the implementation process of the algorithm is simulated; the results of the particle filter tracking algorithm are given.

2. Conveyor connection

First of all, hardware connection, before connection to ensure that the first mechanical arm and the second mechanical arm connection is completed, the first machine is mainly responsible for loading mechanical arm, the second mechanical arm is mainly used to identify after grabbing the object block and placed in the designated position, the conveyor belt motor line connected to the first mechanical arm

of the main control box interface Stepper1. Connect the power cable of the color sensor to the interface GP1/GP2/GP4/GP5 of the second robot arm (this paper takes the fixed interface GP2 as an example). Connect the power cable of the photoelectric sensor to the second robot arm interface GP1/GP4/GP5 (fixed interface GP4 is used as an example). Conveyor connection is completed

2.1 Conveyor working principle

Press and hold the arm unlock button on the arm to move the arm to the first block position. The control panel in DobotStudio will automatically read the current XYZ axis values and fill in the values in the transport demo. Determine the starting position of the transport 1. Check "Suction cups", let the suction cups suck the blocks, press and hold the arm unlock button, and move the blocks to the starting position of the transport on the conveyor. Uncheck Suction cups and place the blocks on the conveyor belt. The control panel in DobotStudio will automatically read the current XYZ values and move the read values to the demo. Confirm the stop position of the conveyor belt. Set the running speed and time of the conveyor belt in the handling program block to control the running distance of the conveyor belt, so that the blocks stay exactly within the gripping range of the sorting robot arm. The second robotic arm point position is debugged by pressing and holding the small arm unlock button to move the robotic arm to the position where the blocks are parked. Fine tuning can be done through DobotStudio so that the suction cup is in the middle of the position right above the blocks. At this point, the control panel in DobotStudio will automatically read the current XYZ axis values and fill the read values into the sorting demo. Check the suction cups, let the suction cups suck the blocks, press and hold the arm unlock button, and move the blocks above the color sensor (not too far from the color sensor, 5mm~10mm is enough). At this point, the control panel in DobotStudio will automatically read the current XYZ axis values and fill in the read values into the sorting demo.

Press and hold the arm unlock button and use the arm to move the block to the desired position. At this point, the control panel in Dobot Studio will automatically read the current XYZ axis values and fill the values into the sorting demo. Once the pointing is done, click "Save" to save the settings. The camera is a 5 megapixel HD industrial lens with a roll-up exposure method and a minimum exposure time of 28us, the lens is a 12mm fixed focal length lens with F2.8-16, and the fill light has 8 levels of brightness adjustment with a maximum illumination of 40,000lux. The package is equipped with more than 40 vision tools and more than 1,000 image operators.

3. Hardware Design

We used the Dobot Magician[1] dynamic target-based robotic uncalibrated vision servo system, the robotic arm, its vision kit and the conveyor belt from Yuejiang Technology Co. After the discussion and research of our team, we finally determined the parameters, selection, functions and hardware circuit design of the competition devices.

The camera is a 5-megapixel high-definition industrial lens with a minimum exposure time of 28us, a 12mm fixed-focus lens with F2.8-16, and a fill light with 8 levels of brightness adjustment and a maximum illumination of 40,000lux. It is easy to use and can easily realize image processing functions such as target 20 matching, defect detection, deformation measurement and character recognition, saving project development time and helping vision recognition projects develop quickly and reliably.

4. Software Design

The software used for programming and debugging the robot is combined with the robot, and the platform can be selected from DobotVisionStudio, DobotStudio, or a freely chosen development environment using the SDK. In terms of software application, we used DobotVisionStudio software as a real-time control interface for the vision device and the robot arm, and whether the entire vision device can recognize the material can be displayed in the management interface of DobotVisionStudio[2] software, while its recognition process and status are also monitored by the software, and the operation of the entire robot arm and vision The operation of the whole robot arm and the vision robot is controlled by the program, and the algorithm analysis of the modified operation program and the control strategy of each link are shown in Figure 1.



Figure 1: DobotVisionStudio terminal interface

4.1 Software design ideas

In the software design, the first need for users to download DobotVisionStudio1.4.2 software, the best book win7 \ win10 system to run the software, after the completion of the download to download python3.7, python3.10 two versions of the programming language, after the need to make up for the need to run the file library, after the need to use the software with a dongle. After entering the software, there are mainly three parts.

The first time you enter, you need to select the general program, create a new program, open the camera configuration, select the camera for the first time, then quickly match the camera parameters, then select the picture storage, you need to create a new folder in which to put the picture named "1" and in the "1" folder as the folder for cropping pictures, and so on to create four new folders, and in each folder under the cropping program for use, and then return to the program to open the save picture, connect the camera and save the picture, and then quickly set the picture storage, select the picture format as "Jpg" format, select the storage method for "src" and then start running the program as the second machine for identification, need to take 1000 pictures of each item at different angles, then crop the pictures to put the object in the center of the picture in order to improve Accuracy, let the machine learn, after learning to judge the accuracy, need to learn more times, and then choose the best.

4.2 Image Acquisition

Double-click the local image to set the parameters, the main parameters are pixel format, fetch interval, scheme save and trigger setting. Pixel format: You can set the pixel format as MONO8 or RGB24. Fetch interval: You can set the fetch interval of auto switch in ms. Save scheme: You can set whether to include the local image when saving the scheme. Auto switch: Each run will switch to the next image. Character trigger filtering: You can control whether the function module runs or not through external communication when it is turned on. Input characters: Select the source of input characters.

4.3 Camera images

Drag the "Camera Image" module to the process editing area, pull down the "Select Camera" column to see all the cameras currently online, and select the camera you want to connect. In software mode, click "single run" to trigger one camera image; click "continuous run" to continuously preview the image, and adjust the parameters to select the camera according to the requirements. You can select GigE, line array camera or U3V camera that is currently online in the LAN to connect. Image Width, Image Height: You can view and set the image width and height of the currently connected camera. Frame Rate: You

can set the frame rate of the connected camera, which affects the speed of image acquisition. Actual frame rate: the real time frame rate of the current camera. Exposure time: The exposure time of the currently opened camera, the exposure affects the brightness of the image. Pixel format: there are two pixel formats, Mono8 and RGB8Ppacked. disconnect reconnect time: when the camera is disconnected due to network and other factors, the module will perform reconnect operation during this time. Gain: Increase the brightness by increasing the gain without increasing the exposure value. Gamma: Gamma correction provides a non-linear mapping mechanism for the output, with Gamma values between 0 and 1, the brightness of dark areas of the image is increased; Gamma [3] values between 1 and 4, the brightness of dark areas of the image is decreased. Line frequency: When the connected camera is a line array camera, the line frequency of the camera can be set. Actual line frequency: the line frequency during the actual operation. Trigger source: You can choose the trigger source according to your needs, including soft trigger for DobotVisionStudio to control the trigger camera, or hard trigger, which needs to be set with the external hardware. Trigger delay: After receiving the trigger signal, the trigger delay is set to generate the corresponding time program. Character trigger filtering: After opening, you can control whether the function module runs through external communication. Input character: select the source of input character.

Trigger character: Any character can trigger the process when the character is not set, and the process can be triggered when the corresponding character is transmitted after setting the character, and the process will not be triggered when the transmitted character is inconsistent with the set character.

Drag the "Store Image" module to the process editing area, double-click to configure the corresponding parameters, run the process after the configuration is completed, and store the camera image, local image or image processing tool processed image, specific parameter configuration trigger to store the image: you can set the trigger variable, with the storage conditions to store the image, the storage conditions are all Save, save when OK, save when NG and no save. Save path: You can modify the path for storing images. Maximum save quantity: The maximum number of images that can be saved under the set path. Storage method: Set the way to process the image when the maximum storage quantity is reached or the disk space is not enough, you can choose two ways to overwrite the previous image or stop storing the image. Save format: There are two types of image formats: BMP and JPEG.

Image naming: you can customize the prefix or subscribe to the previous module data as the prefix and the serial number or date as the suffix, such as IMG-1, and the naming format will change with the module status when using the trigger to store the image, such as IMG-OK-1.

4.4 Feature Matching

Feature matching is divided into high-precision feature matching and fast feature matching. This tool uses the edge features of an image as a template, determines the search space according to the preset parameters, and searches for targets similar to the template in the image, which can be used for locating, counting and determining the presence or absence.

High-precision feature matching has high accuracy and runs slower than fast matching, but the set of features is more refined and the matching accuracy is high. The geometric feature learning model of the template image is used to find and match the target image, which is usually used with the image.

Double-click the feature matching module to configure the parameters, including basic parameters, feature templates, operation parameters and result display. The basic parameters and result display are described in the tool application examples section, and only the feature template and runtime parameters are described here.

4.5 Feature Templates

The shortcut keys in area 4 of the template configuration, from left to right, indicate Move Image, Box Round Modeling Area, Box Rectangular Modeling Area, Customize Polygon Modeling Area with up to 32 vertices, Select Model Matching Center, Generate Model, Erase Outline Point, Clear All Modeling Areas, Undo, and Return. Matching point: Used to create the position reference, you can click "Select Model Matching Center" first, and then select the matching center point in the figure to set by yourself. Scale mode: The principle is that if the automatic mode can meet the demand, no adjustment will be made, and then switch to manual mode if the automatic mode cannot meet the demand. Coarse scale: coarse feature scale parameter, the larger the value, the larger the feature scale, the more sparse the corresponding extracted edge points, but will speed up the feature matching speed. Fine scale: the

fineness of the extracted feature particles, which can only be taken as an integer and is not larger than the rough scale, and is most fine when the value is 1. Threshold mode: The principle is not to adjust if the automatic mode can meet the demand, and then switch to the manual mode if the automatic mode cannot meet the demand. Contrast threshold: This value indicates the size of the contrast, mainly related to the difference between the gray value of the feature points and the surrounding background, the larger the value, the more feature points are eliminated.

4.6 BLOB Analysis

Blob analysis, i.e. the process of detecting, locating or analyzing a target object in an image area whose pixels are of finite gray level, is a tool that provides information about certain characteristics of the target object in the image, such as presence, number, location, shape, orientation, and topological relationships among the blobs. The basic parameters and results display are described in the tool application examples section, and the operational parameters are explained here. When Soft Threshold (Relative) is selected for the Threshold Method: Soft Threshold Softness: parameter for the degree of variation between Soft Threshold Low and High Threshold. Soft Threshold Low Tail: The left tail of the threshold, which can be removed by percentage example. Soft Threshold High Tail: the right tail of the threshold, which can be removed by percentage example. Low threshold proportion: the proportion of low thresholds in the soft threshold range. High threshold proportion: the proportion of high thresholds within the soft threshold range. Polarity: there are two modes: darker-than-background and brighter-than-background. Darker-than-background is the feature image pixel value lower than the background pixel value; brighter-than-background is the feature image pixel value higher than the background pixel value. Threshold range: the threshold value of the feature calculation for filtering the target blob, including low threshold and high threshold[4], a review of the application of deep learning in the field of image processing, the calculated value of the blob feature is judged as the target blob if it is within the threshold range, otherwise it is evaluated as non-target blob and is not output. Number of lookups: Set the number of lookups for Blob graphics.

Minimum area of hole: The minimum area of non-blob area that the content of the blob area can tolerate, if it is not greater than this value, the hole will be filled as a blob. Enable: If the current feature enable is on, the feature will be used for blob filtering, if it is off, the feature will not be used for blob filtering.

4.7 Deep Learning

Deep learning is a class of machine learning algorithms developed from traditional neural networks. Through deep learning models similar to brain neural networks, computers can absorb, learn and understand complex information in the real world like humans, and accomplish difficult recognition tasks, which can be used for character localization, recognition, defect detection, image classification, and target detection.

Deep learning is based on a certain data base, so a large number of data sets need to be trained before deep learning, and the data sets participating in the training need to be labeled to ensure the diversity of data as much as possible.

This function module requires high computer configuration and does not support RTX series graphics cards. Deep learning classification, target detection, text location, defect detection, and character recognition modules require more than 2G video memory, and need to be paired with[5] Opportunities, challenges, and practical paths for the integration of artificial intelligence and journal development, and Geforce GTX960 or above video cards. Deep learning training tool video memory requires 4G and above, and needs to be paired with GTX 1080 and above graphics cards. Gigabyte GTX1080Ti, Asus GTX1080Ti, NVIDIA GTX1080Ti graphics cards are recommended.

DL character localization, classification and target detection are available in two ways, G and C. The module name of GPU version is suffixed with "G", and the module name of CPU version is suffixed with "C", CPU version is not dependent on graphics card. This function needs to be used with a special dongle for deep learning.

4.8 Calibration plate calibration

The calibration board is divided into two types of calibration boards: checkerboard and circle. The software will calculate the mapping matrix between the image coordinate system and the physical

coordinate system of the checkerboard, the calibration error, the calibration status, and click Generate Calibration File to complete the calibration. The tool will generate a calibration file for calibration conversion. The Generate Calibration File button allows you to select the path to save the generated calibration file. Origin (x), Origin (y): The origin is the origin of the physical coordinates, you can set the coordinates of the origin, i.e. the position of the origin of the X and Y axes in the figure.

Rotation angle: the rotation angle of the calibration plate.

Coordinate system mode: choose left hand coordinate system or right hand coordinate system.

Physical size: The length of each black and white cell of the checkerboard grid or the distance between the two adjacent centers of the board in mm.

Calibration board type: divided into checkerboard grid calibration board and circle calibration board.

Degrees of freedom: There are three types of scaling, rotation, aspect ratio, tilt, translation and perspective, scaling, rotation, aspect ratio, tilt and translation, scaling, rotation and translation, and the three parameter settings correspond to "perspective transformation", "affine transformation" and "similarity transformation" respectively. The three parameter settings correspond to "Perspective Transformation", "Affine Transformation" and "Similarity Transformation". Grayscale Contrast: The minimum value of contrast between adjacent black and white cells of the checkerboard image, it is recommended to use the default value. Median filter state: whether to perform median filtering before extracting corner points, there are two modes: "Perform filtering" and "No filtering", we recommend using the default value. Sub-pixel window: This parameter indicates whether the window size is adaptive to calculate the sub-pixel accuracy of the corner points, when each square of the checkerboard grid occupies more pixels, this value can be increased appropriately. Weight function: The least squares, Huber, and Tukey algorithm functions can be selected. It is recommended to use the default parameter settings. The weight factor is the clipping factor of the corresponding method, and it is recommended to use the default value.

5. System development and debugging

The loading robot loads the material modules to be sorted onto the conveyor belt. The material modules stop when they reach the bottom of the camera and are identified by AI as the type of material. The whole process needs to be controlled by a Python program. Based on the original accompanying operating program, only the loading robot can pick up the material and load it onto the conveyor belt, then the conveyor belt moves and stops when the material is transported to the photoelectric switch identification area. The sorting robot operates, picks up the material and places it in the sorting area. The following is a simple system using the Python language

First open python into the main program, create a TCP client model inference, before inference, be sure to call `model.eval()` to set dropout and batch normalization layer for evaluation mode after OpenCV to PIL format PIL image data conversion to tensor data, and normalized image pre-processing, image processing: grayscale, image pre-processing, image pre-processing. image processing: grayscale, image processing: binarization, image processing: morphology open operation, image processing: envelope contour if the contour is not found, return to the original image to prevent the error reported by the empty load, print the contour area, if the extracted contour area is less than 300000, skip, continue to find the contour, get all contours (Blob connected area) the smallest outer rectangle (center (x,y), (width,height), rotation angle) data object `rect`, get the coordinates of the four vertices of the smallest external rectangle (top left, bottom left, bottom right, top right), data type conversion, according to the center coordinates of the smallest external rectangle and angle, build a rotation matrix `rot_img`

input the target outline rectangle center coordinates `center = rect[0]`, rectangle angle = use the four rectangle vertex coordinates array `[box]` obtained earlier, draw the outline in the original layer using the rotation matrix `rot_img`, the original layer `img` to achieve the center affine transformation, the size of the transformed layer remains unchanged after the affine transformation to get `img_waf` layer center point coordinates and the target outline minimum external rectangle center point, according to the data object `rect` extract the rectangle center point coordinates and width and height `x, y, w, h`, respectively, to get the rectangle row (x) and column (y) start and end points in `img_waf` cut out the target layer after correcting the angle `src_roi` display inference results, set the picture display resolution, draw text information, get real-time time as named, save the recognition picture and results to the folder cropping coordinates `[y0:y1, x0:x1]` ROI area (can be modified according to user requirements) save the image, clear the cache picture, determine whether there is a result folder, traverse the path of the stitching file, determine whether the

path object is a file, delete the file, create a new result folder, enter the main program entry, clear the last run recognition result cache image before each run, ROI crop the original image, preprocess the cropped image, reason the image preprocessed image to the model, print the reasoning result, display the reasoning result

6. Conclusion

Through visual robotics item awareness. This aspect of vision robotics has provided a greater awareness, which has not only facilitated the understanding of vision robotics artificial intelligence, but has also broadened the horizon. However, after participating in this vision robotics project, I realized that robots can be specifically subdivided into many modules, and that the vision robotics module is promising enough for development. In traditional applications, industrial robots are mostly programmed offline or taught to work in a strictly functional way, which does not meet the requirements of flexible manufacturing. The introduction of machine vision technology has improved the flexibility of robots in performing tasks. The initial and final positions of the robot can be calculated adaptively from the vision acquisition data, and with the collaboration between the robot and the vision system, not only can the robot's motion be used to expand the visual range of the camera, but also the camera can be used to assist the robot in positioning and completing complex tasks that are difficult to achieve manually. This will improve the intelligence of industrial robots. Through study and review, we have concluded that vision robots have more advantages and potential for development than traditional robots. It is believed that vision robots will be widely used in the future.

References

- [1] Zhao Xifeng, Du Xiangdang, Gong Jingjing, et al. Simulation of a dynamic target-based robot calibration-free vision servo system[J]. *Machine Tools and Hydraulics*, 2012, 40 (11) :116-119.
- [2] Yang ZM, Liu JX, Feng XY, et al. Preparation and mechanical properties of basalt fiber cloth reinforced composite materials with different resin matrix [J]. *Journal of Beijing University of Technology*, 2018, 38 (6): 655-660.
- [3] Luo Xiaohui. Machine learning behind artificial intelligence [J]. *Electronic World*. 2019 (14)
- [4] Yin Qilin, Wang Jinwei. A review of deep learning in image processing[J]. *Journal of Higher Education*. 2018(09)
- [5] Chen H, Liu Yumeng, Pei Meng. Research on the opportunities, challenges and practical paths for the integration of artificial intelligence and journal development[J]. *Research on Chinese science and technology journals*. 2019(03).