

## Vision

### 1.1.

The software structure now is based on ROS framework, which greatly increases the efficiency of multitasking coordination and simplifies the process of application of new algorithm.

There are four modules of the algorithm of our robot, vision is one part of them. Vision Module is used to recognize objects in the football court and convert them into information in geometric forms. The vision part mainly deal with two things: object recognition and distance measurement. We use one single camera, and our vision algorithm is based on computer vision (OpenCV). This year we use deep learning methods to increase the accuracy of the recognition of the ball and the sideline, and calculate the distance. Then the data will be used to decide the robot's behavior.

### 1.2. The application of deep learning in ball searching

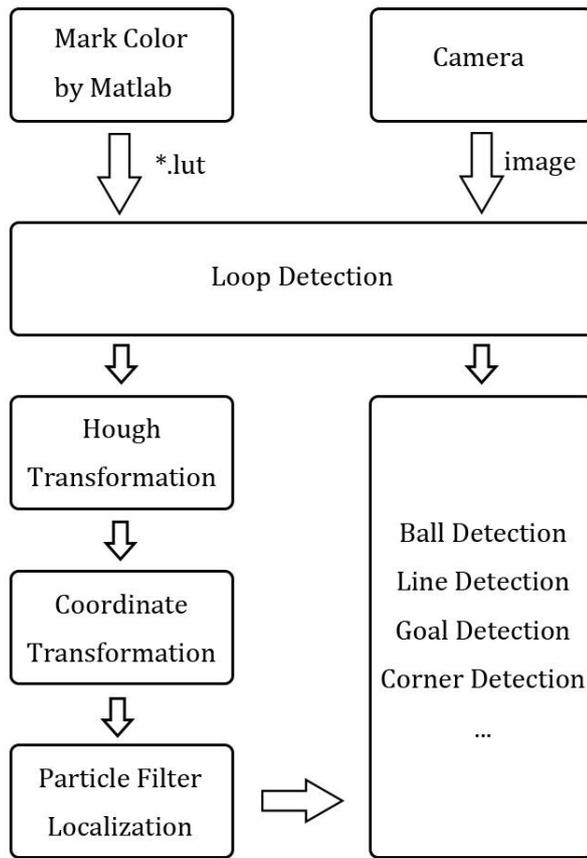
The task of this part is to quickly recognize the low-resolution image and return a vector representing the position and distance of the ball. The motion mechanism looks for the ball in the direction indicated by the vector.

Our artificial neural network system adopts nonlinear decision making, which is constructed by two layers of sigmoid elements, that is, one layer of hidden layer and one layer of output layer. The size of the hidden layer is determined according to the actual calculation ability and the technical details of the processor. The size of the output layer is no less than 3, and the relative position of the ball is determined according to the output vector.

In the aspect of input coding, we adopt gray-scale maps of 32p resolution which has been preprocessed. In order to increase the operation speed, the gray scale of each pixel is obtained by random sampling of the corresponding region of the source image. This strategy makes it easy to convert input into vector processing.

In the aspect of output coding, we combine direction and distance in the vector to indicate the position of the ball. In order to accelerate the training process, the minimum value of the standard output of the training sample is slightly greater than 0, and the maximum value is slightly less than 1.

The neural network structure adopts the standard three-layer acyclic sigmoid network. In the process of training, the neural network modifies the weight vector while cross-verifying the size of the hidden layer. That is to say, lack of hidden layers will lead to poor fitting effect and waste of training samples. In contrast, too many hidden layers will lead to over-fitting and high training cost. We will add impulse term to enhance accuracy and speed up training in the process of applying backpropagation algorithm.



1. O. Chutatape, Linfeng Guo. A modified Hough transform for line detection and its performance. Pattern Recognition, 32(2): 181-192, 1999.
2. S. Thrun, W. Burgard, D. Fox. Probabilistic Robotics. MIT press, 2005