Autonomous Underwater Vehicles for the RoboSub Competition

Wenzhi Liu, Haibo Li, Lin Zhang, Xiaoyang Wang, Hanwen Zhou , Hong Liu,Xiangren Sun,Jingqi Wang

***Abstract*—The under water robot team of HEU(Harbin Engineering University) has been built for many years. In recent years we have attended many robot competitions, and gained many experiences. We constantly improve our robot and hope our robot to have perfect performance . An important improvement upon the previous system was to use Convolution neural networking image processing.**

I. INTRODUCTION

Underwater system has to face many challenges such as disquieted communication, uncertain control, independent navigation. Robosub provides a competition platform that encourage robotic funs overcome these difficulties and promote the development of autonomous underwater techniques.

HEU-AUV team's goal is to research and implement a smart robot that generally handle underwater operations. Ao ming which is for research and experiment purpose is designed to reach our goal. Special facilities and functionalities are given to successfully complete the challenging courses in Robosub Competition, held in San Diego annually on July. The courses such as firing torpedo at small target, locating sonar pinger and dropping markers into a bin etc. must be completed automatically without any human interference. To this goal Aoming needs to be carefully designed so that it has precise control, sensitive sensors and being robust enough in target recognition.

Our team is divided into visual, software, hardware, mechanical group, to achieve greater efficiency and better results.

II. DESIGN OVERVIEW

Robot is designed for streamlined, so the robot can get a high movement speed, robot achieve 5 degrees of motions though five brushless motor control.

Brushless motor arrangement makes the thrust is not coupled with each other. 2 sets of independent 5000mah Ni-MH battery supply power for motors and other part like PC and DVL.

Robot equipped with ultra-short baseline hydrophone to achieve acoustic navigation, DVL and INS to get robot's relatively location.

The computer uses the NVIDIA platform Jetson TX1, the program runs under the ROS software framework, the front view and the stereo camera below robot body to help obtain information like the location of the target to complete the visual servo task.

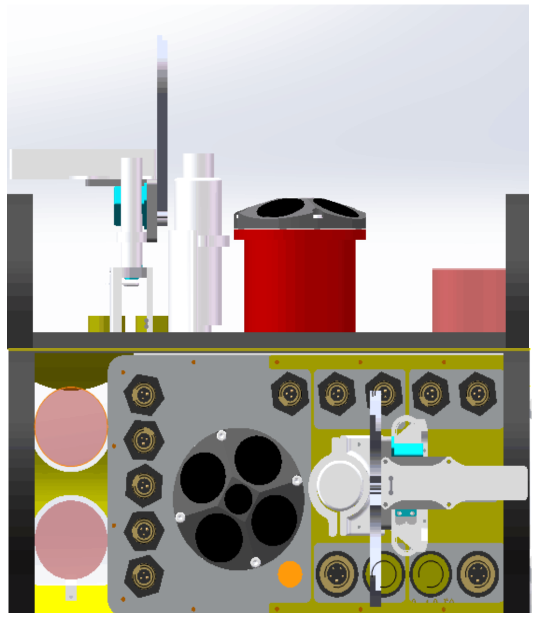

Fig 1. Marker Delivery device structure diagram

Hydrophones placed in the 4 diagonal is mainly for positioning of underwater beacon through the correlation algorithm.

Fig 2. Marker Delivery device structure diagram

III. MECHANICAL SYSTEMS.

A. The Hull

Early in Ao Ming's design process, we decided to adopt module design method, which means each module of Ao Ming can be removed and installed individually, so that the faulty hardware can be easily replaced. More than this, the module design makes Ao Ming easier to assemble or transport.

Before manufacturing, the components of Ao Ming are designed with SolidWorks. After all of the components have been designed, the ANSYS is used to do finite element analysis, to make sure that the buoyancy of Ao Ming is slightly more than it's gravity.

The hull was firstly printed by 3D printing technology, then it was greatly reinforced by carbon fiber material to make it light and strong. The position holes of AUV is processed by CNC to ensure its accuracy and extend its service life.

Ao Ming's mechanical system consists of three parts, including the nose, the middle ship, and the stern. The large hull provides installation points and protection for components, sensors and actuators. Ao Ming's mechanical system uses a bionic streamlined design which imitates the marine fish, and the structure of Ao Ming is similar to modern submarines. Therefore, the structure of the AUV can reduce the resistance and adapt to the underwater fluid environments perfectly. There are two windows on the upper part of the hull for installing the counter weights. At the same time, a pair of horizontal tail fins are added to make the AUV able to sail straightly.

B. The Design of the Power Module

Ao Ming uses five powerful brushless motors in total, including 2 vertical thrusters, 2 lateral thrusters, and 1 main thruster. This design of motor distribution makes the thrust from each direction uncoupled, so that the AUV can be more controllable. With the help of the 5 thrusters, Ao Ming can move freely in 5 degrees, meet the requirements of the task.

The planetary deceleration mechanism of the propellers gives Ao Ming a powerful thrust, and using highly efficient blades makes Ao Ming agile. Four of the propellers are loaded into ducts, which can not only protect the propellers, but also make the motor provide stronger thrust. What's more, we chose new ESCs for this year, which can enhance the low-speed performance of the motors to complete the precise control of the robot.

C. External Structure

Ao Ming's external structure includes main cabin, hydrophone, DVL, cameras, manipulator, Torpedo launching mechanism, and throwing mechanism.

The main cabin is placed on the upper side of the middle ship, and it's large enough to hold all the circuits and computers. The equipment in the main cabin are connected with outside through the watertight plugs to ensure the tightness of the main cabin.

The battery cabin, which is made of a acrylic cylier, is placed close to the main cabin, and it also has its own-designed watertight structure.

The DVL and hydrophone are fixed on the main cabin, just under the center of AUV.

The forward-looking camera is fixed on the nose to AUV, and the look-down camera, manipulator, and throwing mechanism are closely fixed on the bottom of the main cain. That design saves space and enables the convenient visual control.

The manipulator of Ao Ming is an unidirection mechanical gripper, which is controlled by a actuator. It grips the handle simply and reliably by opening and closing.

The torpedo launching mechanism was designed by ejection method. While the torpedo needs to be ejected, the trigger mechanism will trigger the movement of the spring and eject the torpedo.

Before being thrown, the projectile is attached to the bottom of the robot with magnetic. When the projectile needs to be thrown, a servo motor will be used to trigger the release mechanism, so that the projectile is free to sink. In order to ensure that the projectile' s trajectory is vertical, the shape and weight of the projectile are carefully designed.

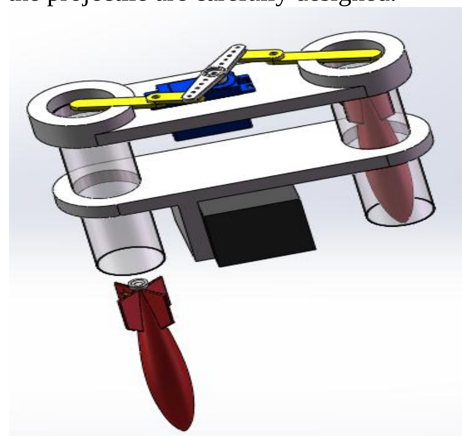

Fig 3. Marker Delivery device structure diagram

IV. The Electrical System

A. Power management and underlying communication modules

The module as Fig 4.is responsible for all hardware and power supply management of the sensor, the use of sampling resistor with high input common mode voltage op amp to obtain the size of the current, and the voltage detection. Power management module with a number of regulator module to complete the supply of multiple parts. There is an embedded PC -GM45 with P7350 CPU in our vehicle to do image processing work and navigation. There is a STM32 control board in our vehicle to control the motors. All the data of sensors are processed by STC12C5A60S2. A DSP board is used to process sonar information.

At the same time the module integrates CAN interface and serial port, it through the serial port and CAN interface to collect all the sensor data, the collected data sent to the Jetson Tx1 serial port, and through the serial port to receive Jeton TX1 data sent to the motor through the CAN drive module.

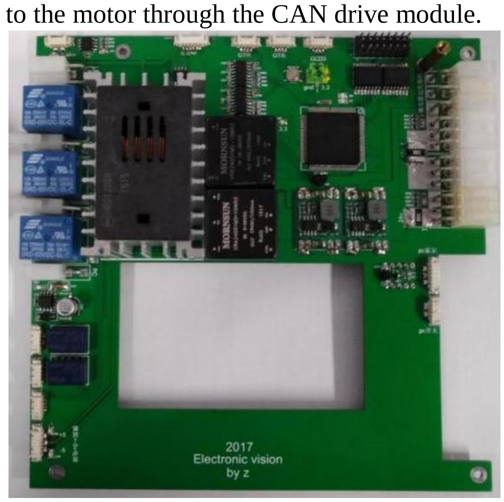

Fig4. Power management system

B. Motor drive module

Motor drive module as Fig 5 with CAN bus interface, receive A module to send the data control Brushless motor, the board with Hall switch control relay, with emergency stop switch can be an emergency shutdown motor.

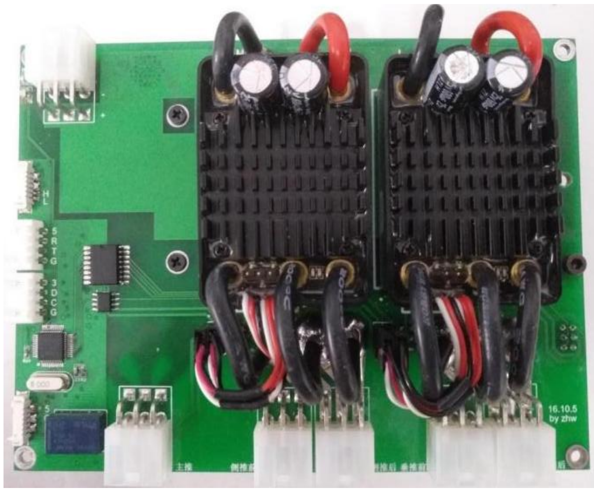

Fig5. Drive module

C. Sensor

The hydrophone signal processing system as Fig 6 comprises two parts, a signal conditioning part and a signal processing part.

Signal conditioning part of the three-pass band-pass filter amplifier circuit, the original signal from the transducer back through the three-way fourth-order Butterworth filter processing, band-pass filter bandwidth of 18k to 42kHZ, magnification of 52db, The Butterworth filter is implemented using the integrated filter chip MAX274, which converts the analog signal to the hydrophone signal processor via a shielded wire.

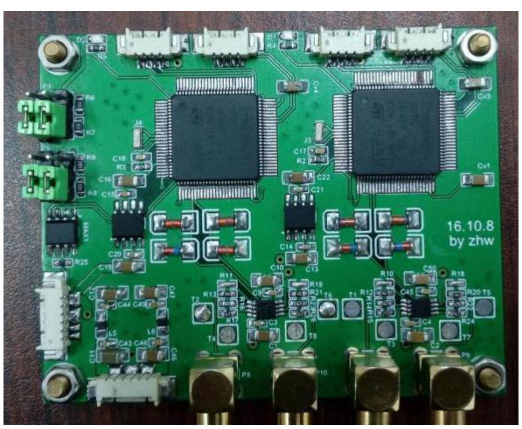

Fig 6: hydrophone signal processing system

In the calculation of dft, if the length of the discrete periodic sinusoidal signal is not an integer multiple of the period of the measured signal, a spectral leakage occurs, resulting in a phase calculation error. And because fft can only take a fixed number of points, usually do not meet this requirement, so we use dft to calculate the phase, fft or dft to calculate the frequency.

As the stm32 speed limit, run the c language to achieve the dft will spend a lot of time in the signal to determine whether the signal to be the corresponding phase, the use of fft, calculate the phase phase using dft, CAN bus communication data back.

Depth sensor using MEMS depth sensor chip, the range is 0 ~ 10bar, sealed in the waterproof plug inside, the data CAN bus back.

The attitude sensor uses MPU9250, the main reading angular velocity information to make up the fiber optic gyro can not output the angular velocity problem. The angular velocity of the output is mainly for more angular velocity observations and better control performance. If the angular differential is used, the high frequency noise is amplified. Using CAN bus back.

The relationship between the various hardware subsystems is shown in Fig7.

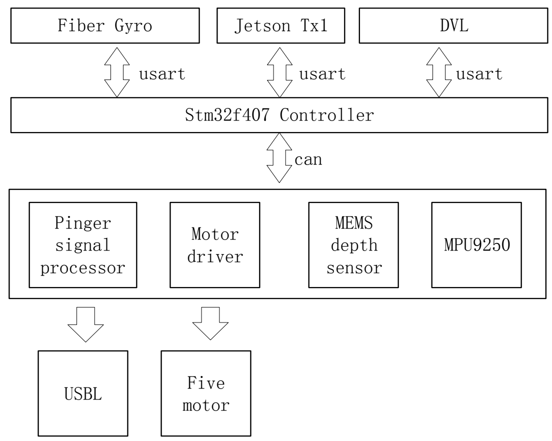

Fig7. hardware framework

V. Software

The software is constituted with five main units: Vision target detection, motion control, stereo odometry, adaptive vision servo and multi-sensor fused navigation. The connections and interactions among these units are described in Fig8. All units are constructed upon Ubuntu 16.04 and ROS Kinatic both of which are installed on Nvidia Jetson TX1 Embedded Kits.

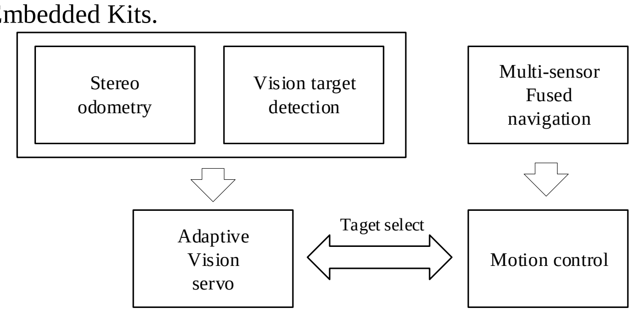

Fig 8:software framework

A. Target detection

This unit is powered by a 15-layer convolutional neural network. The raw RGB image is feed into the network. Behind several conv-and-pooling operations is an operation of reshape, which flats a 64 channel 2D feature image to form a vector. Then, the vector is feed into two fully-connected layers and is mapped into a vector. This vector is then reshaped to a 10-channel 7*7 grid maps. That means, each grid contains a 10-dimision descriptor which contains detailed detecting and classification results. The first four elements describes a point (x,y) in frame which is the location of the left upper corner of target bounding box. The third and fourth describes width and length of the bounding box respectively. The fifth element describes the probability of the existence of a target in respective grid areas of the frame. We train the network with samples of operating targets of the competition .We collect the real target's image as Fig 9 from the webside.And Conversion to standard VOC format.

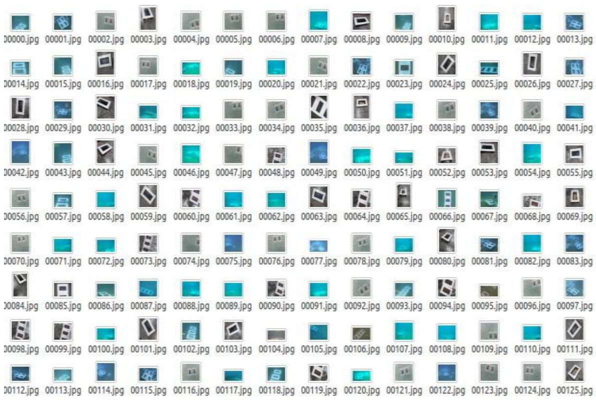

Fig 9:image of tagets

The pre-train model is trained by voc 2007. In each training iteration a batch of 64 images and their respective ground-truth targets (bounding box that contains target in raw frame and the classification results) are extracted from the training data set. After training 20000 iterations we test the model with the test set. Which contains about 30% of total image samples in the dataset. The modified (trained) model is capable enough to tell the control and servo system what it has seen and where is it. Fig10 illustrates the detecting information generated by deep convolutional neural-network and get the location of box by stereo camera at the same time. Fig11 is disparity map. All processing units runs synchronically on Nvidia GPU architecture. That saves computation time assumption by 70 times.

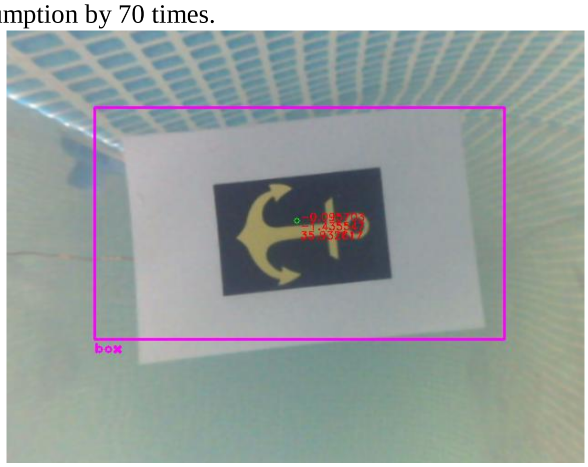

Fig 10. YOLO and stereo camera's result

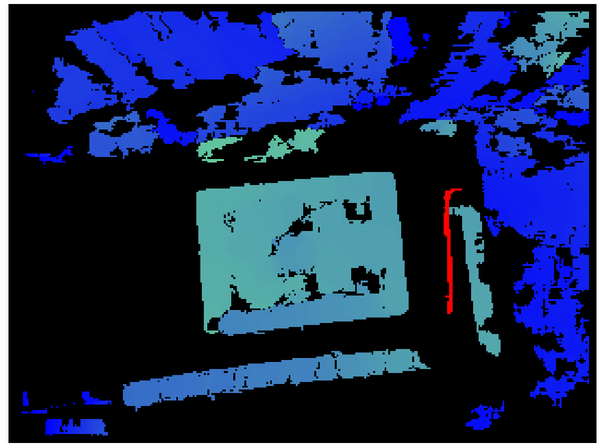

Fig 11. disparity map

B. Motion control

The motion control unit acts on four degree of freedoms: Surge, Sway, Heave and Yaw. It receives a point in earth fixed frame, namely the reference pose. The pose and altitude of the robot which is measured by navigation system is transmitted to the control unit as a feedback. Simply subtract the measured pose from the reference pose to derive the error vector in fixed frame. Then, a transformation matrix is calculated based on the information of altitude. This transformation matrix projects the error vector in fixed frame into body frame which is parallel with the thruster forces. The error in body frame is feed to four PID controllers to generate motor commands, these controllers stabilize the four degrees of freedoms respectively. The amplitude of the respective motor command is limited in the range of [-110 110]. The framework of the control unit is described in Fig12. We implement a derivative prioritized version of PID approach to promote the performance in condition of dynamic reference frame.

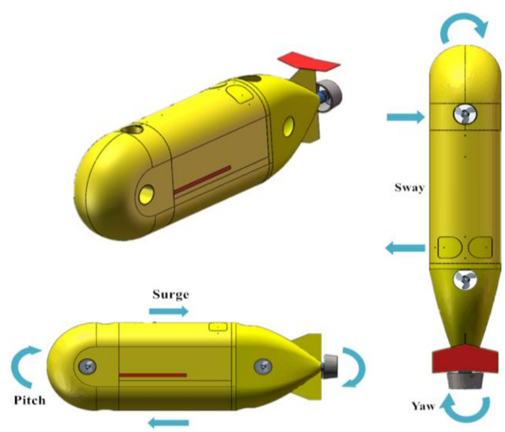

Fig 12. framework of the control unit

C. Stereo Odometry

We apply a block stereo matching in term of the detected bounding boxes. The distance between camera and the target is derived by averaging over a deterministic region around the center of bounding box. During the experiment, the center of bounding box suffers inaccuracy caused by the video signal disturbance and computational delay. We address this problem by applying the watershed image segmentation to extract the contours of regions with specific colors. The segmentation algorithm process the region of interest (ROI) defined by formal deep learning detection, derive the contours and derive the center of the contours by computing the moment of 2D geometry object. The segment processing is much faster, it overcomes the influence of video signal disturbance with the operations of corrosion and dilating. Stereo matching thread runs in parallel with classification thread which makes sure that the sensor messages are processed on time.

D. The vision servo policy optimization

After the target center 3D position in camera frame has been derived, we use a PID controller guiding policy search algorithm for searching the optimal or strictly, near-optimal vision servo policy which maximum the probability of successful vision operation. The policy is modeled by a neural-network whose update is driven by numerical analyzing. This part of technology belongs to the framework of reinforcement learning and is an implementation of stochastic policy gradient (SPG). We use python for developing computation graph of machine learning. The optimization target is defined as the accumulated rewards of state-action pairs. The technique of importance sampling is applied to predict the accumulated discounted rewards of the neural-network police under the trajectory generated by PID controllers. Since the moving speed of the robot is unknown, which means the agent search the optimal policy in partial observable environment, we add LSTM layer before the FC layers to make it remember the past and observe the speed indirectly with state transformation trend. The PID vision servo controller is used so that the training speed obviously promoted. After training, the neural network policy even outperforms its PID demonstrator in vision tasks.

Basic algorithms are constructed using C++ and are with the assist of OpenCV and ROS. Besides, deep neural-network learning is constructed with the assist of Google TensorFlow. Each unit we illustrated in this section are built in a functionality package. They communicate with each other with ROS topics and services.

VI. References

[1] Redmon, Joseph, et al. "You Only Look Once: Unified, Real-Time Object Detection." Computer Vision and Pattern Recognition IEEE, 2016:779-788.

[2] Simonyan, Karen, and A. Zisserman. "Very Deep Convolutional Networks for Large-Scale Image Recognition." Computer Science (2014).

[3] Sermanet, Pierre, et al. "OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks." Eprint Arxiv (2013).

[4] Szegedy, Christian, et al. "Going deeper with convolutions." Computer Vision and Pattern Recognition IEEE, 2015:1-9.

[5] Girshick, Ross. "2015 IEEE International Conference on Computer Vision (ICCV)." IEEE International Conference on Computer Vision IEEE, 2015:1440-1448.

Appendix

The underwater robot team of HEU (Harbin Engineering University) has been built for many years. And our university is located in the beautiful city, Harbin. It was founded in 1953.The Department of Naval Architecture and Ocean Engineering (NAOE) at the College of Shipbuilding Engineering has a history of more than 50 years. Originally established as the Naval Engineering Department in 1953, it was one of the key departments of Harbin Military Engineering Institute. Our team rely on the school's scientific research resources and focus on the development of the underwater vehicle. In recent years we have attended many robot competitions and gained many experiences. At the same time, we constantly improve our robot and hope the robot perfect well in practice. In order to improve security and flexibility of the robot, the team design the vehicle though various aspects, especially the design of the software. This year, we will take part in the competition with our new work. We want to communicate with other teams during through this competition. We sincerely welcome the other team to come to our school and exchange ideas of the robot during the long-term development.

Last year, our team get the sixth good results. This year, we made more full preparation and hoped get better ranking. Simultaneously, we will learn more to improve out technical level.