Proceedings of the 40th Chinese Control Conference
July 26-28, 2021, Shanghai, China

Data Calibration Algorithm for Artificial Lateral Line Sensor of Robotic Fish on Improved LSTM

Yaming Ou[1,2], Zhuoliang Zhang[2], Chao Zhou[3], Bo Zhou[1]
1. School Of Automation, Southeast University, Nanjing 210096, China, E-mail: kingo.sunshine@gmail.com
2. State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
3. Naval Research Academy, Beijing 100072, China

Abstract: All the time, the measurement of swimming speed has long been a major research challenge in the field of robotic fish. The artificial lateral line system can accurately measure the speed without modifying the structure of the robot fish too much. In this paper, we developed a cantilever artificial lateral line(ALL) sensor based on piezoresistive effect by imitating the "SNs" cell structure of real fish, and established an accurate physical model for it considering the complex underwater boundary layer effect, which is conducive to the selection of hardware parameters such as sensor size and stiffness. In addition, a flowmeter-based speed calibration experimental platform was developed, based on which we proposed an "end to end" speed calibration algorithm based LSTM without filtering. Then a potential mapping relationship between the electrical signal of the sensor and the speed value of the flowmeter was obtained, which laid a strong foundation for the subsequent measurement of free-swimming speed of robotic fish by artificial lateral line system composed of this sensor.

Key Words: velocity measurement, robotic fish, artificial lateral line sensor, data processing

1 Introduction

In recent years, bionic robot is a hot research topic in the field of robot. Robotic fish, as the main research direction in the field of underwater bionics, have been able to realize some complex motor functions, such as fixed height control and 3D motion. However, for the complex underwater environment, the visual environment underwater is very bad due to the absorption of light by the water body and the refraction of light by the particles in the water. In addition, due to the fluctuation of the water itself, it is difficult to measure the pose error when the robot moves, which brings great challenges to the speed measurement of robotic fish. The existing underwater velocity measurement schemes mainly include Doppler meter velocity measurement, inertial navigation system, GPS, etc., which are either large in size or still in the laboratory environment, so it is difficult to be really used in the speed measurement of robotic fish.

Ichthyida is the largest class in the subphylum of vertebrates extant, and has evolved into a variety of fish with complex body shapes.There is no doubt that over the course of long evolution, fish have developed sensory organs that are well adapted to the complex and harsh underwater environment.Stenon[1] found the existence of lateral line systems in fish as early as the 17th century, and Hofer[2] later found that lateral lines in fish sense water flow.Studies have shown that the sensory unit of the lateral line in fish is the colliculus, a receptor organ made up of sensory hair cells and support cells covered by colloidal suckers.Depending on where it is located, there are two types of nerves: superficial Neuromasts (SNs) that lie on the surface of the fish or grow above the skin, and Canal Neuromasts (CNs) that lie inside the fluid-filled tube beneath the skin.Among them, SNs is mainly sensitive to local flow velocity, while CNs is mainly sensitive to flow acceleration.In order to learn from the speed perception of real fish, researchers use the knowledge of bionics to build an artificial lateral line system.

Based on the piezoelectric effect, Asadnia et al.[3] developed a piezoelectric sensor using Pb(Zr0.52Ti0.48)03 as the main functional material, using MEMS micromachining technology and sol-gel method. The sensor requires no power supply, is light in weight, and is easy to mount on the surface. However, its sensitivity at higher frequencies is not high, and the charge is easy to leak, so the safety can not be guaranteed. Dagamseh et al.[4, 5] drew inspiration from the fluid sensitivity of cricket hair and the lateral line system of fish. Combining their advantages, the group used a polysilicon surface micromachining technology to create a capacitive sensor. The sensor has low power consumption, but its load capacity is poor and its output characteristics are nonlinear. Liu et al.[6] proposed a new type of micromechanical hot film side line sensor based on hot-wire anemometer principle (HWA), which was realized through thin film deposition process combined with standard printed circuit. The sensing element is Cr/Ni/Pt, and the resistance temperature coefficient is 2000ppm/K. The flexible printed circuit board is preprinted by magnetron sputtering or pulsed laser deposition. The sensor has good mechanical properties and sensing ability, simple structure, has a certain flexibility, easy to attach to the surface of objects. But this sensor is very low sensitivity, not suitable for accurate speed measurement. In addition, some researchers use off-the-shelf pressure sensors to develop lateral sensors based on the principle of pressure difference. Wang et al.[7] took advantage of the existing pressure sensor Consensic CPS131 on the market and arranged 11 such sensors into an array to form a side line system. On this basis, IMU was used as an auxiliary device to eliminate noise, so as to achieve the purpose of speed measurement. Obviously, although this method can achieve the purpose of velocity measurement, the pressure principle is not the speed sensing principle of fish in nature, but can only be regarded as a "pseudo-artificial lateral line system".

*This work was supported by National Nature Science Foundation of China (62033013, 61903362, 62003341, 61973303).

In addition, in the development process of ALL system, another urgent problem to be solved is how to map the acquired sideline sensor data to the real velocity value, that is, data calibration. At present, most ALL systems are difficult to establish accurate models because of the complex underwater environment, and most of them adopt simple calibration means, resulting in very large errors. Xu et al.[8] used the least square method to fit the relationship equation between the value of the lateral sensor and the velocity, starting from the simulation and the real environment. It can be seen that the error is small in the simulation environment, but the calibration error is large in the complex real underwater environment. Wang et al.[9] derived the quadratic relationship of stagnation pressure, free flow static pressure and free flow pressure by Bernoulli's formula, and directly fitted the relationship between the lateral sensor and the velocity with a second-order polynomial, and the final calibration curve was also greatly different from the real curve.

In order to solve the above-mentioned problems, a cantilever artificial lateral sensor based on piezoresistive effect is proposed in this paper, and the corresponding data calibration method is developed to control the calibration error within a very small range. The lateral line sensor monomer - flow meter calibration system is built, which is also applicable to the calibration experiment of different lateral line sensor monomer; Considering the boundary layer effect in hydrodynamics, an accurate mechanism model of the sensor is established, which provides a theoretical basis for material selection; Based on the idea of sequence feature prediction, an "end-to-end" calibration algorithm based on improved LSTM[10] is developed to complete the data calibration without filters, which effectively avoids complex noise modeling and is very useful for the research of underwater environment.

2 Artificial Lateral Line Sensor

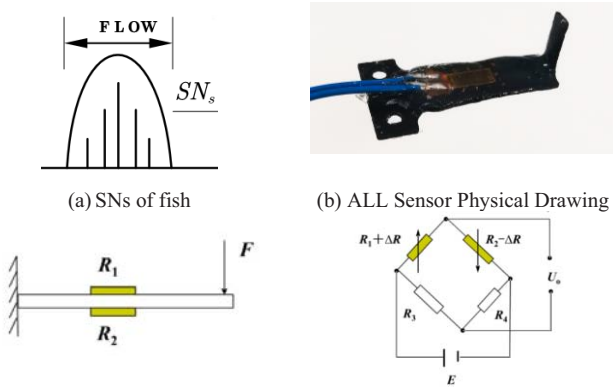
2.1 Physical Design

SNs are located on the surface of the fish and are sensitive to the local flow velocity. As shown in Fig. 1(a), when water strikes the sensory top of SNs, the sensory hair cells located in the lower part release nerve impulses, which are then transmitted to the nerve center, and thus the fish perceives the water flow rate[2].

In order to simulate the lateral line receptor structure of the real fish mentioned above and to reduce the influence of the artificial lateral line on the morphology of the robotic fish, an ALL sensor with a small size as shown in Fig. 1(b) is designed in this paper.

The sensor consists of three parts: horizontal cantilever, headstream plane, and strain gage. Among them, the first two materials are Future 8200 resin, the latter model is BX120-3AA. The horizontal cantilever and the headstream plane together constitute the sensory top structure of a real fish, and the strain gauge is responsible for measuring the deformation of the structure.

In practice, the fixed end of the horizontal cantilever of the sensor is fixed on the surface of the robotic fish. When the headstream plane of the sensor is impacted by the current, the strain gage will strain and its resistance will change



(a) SNs of fish (b) ALL Sensor Physical Drawing

(c) Installation of strain gauge for differential measurement of half bridge (d) Differential half bridge adjacent arm measuring circuit

Fig. 1: Comparison of real fish and ALL sensor

accordingly. The change will be transmitted through the half-bridge circuit in the form of electrical signals.

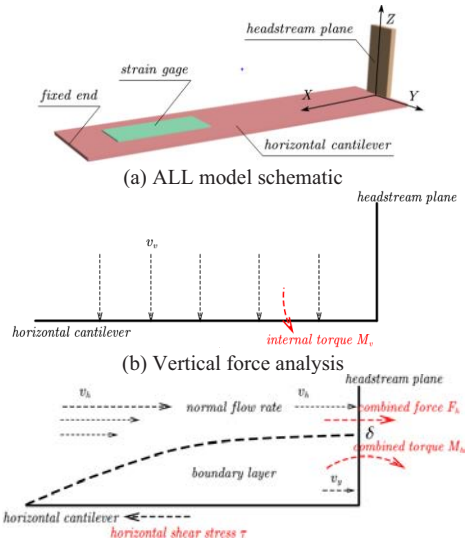
2.2 Mechanistic modeling

Assume that the water flows to the sensor with velocity v and angle θ. As shown in Fig. 2, the force is divided into vertical and horizontal directions, then:

$$v_h = v\cos\theta \qquad v_v = v\sin\theta \tag{1}$$

where v_h represents the partial velocity of the water in the horizontal direction and v_v represents the partial velocity of the water in the vertical direction.

Set the horizontal cantilever length l_1, width b_1, thickness h_1, the headstream plane length l_2, width b_2, thickness h_2.



(a) ALL model schematic

(b) Vertical force analysis

(c) Horizontal force analysis

Fig. 2: ALL force analysis

First, in the vertical direction, as shown in Fig. 2(b), the water flow acts on the horizontal cantilever to produce a bending tension. From the bypass resistance equation

$$F = \frac{1}{2} C_D \rho v^2 A \tag{2}$$

Where C_D is the coefficient of drag around the flow, A is the head-on area, ρ is the fluid density, and v is the flow velocity.

Also, vertical internal moment

$$M_v = \int_0^{l_1} q_v x dx = \int_0^{l_1} \frac{F_v}{l_1} x dx \tag{3}$$

Where, q_v is the linear density under force, and F_v is the resistance around the flow in the vertical direction.

The relationship between the combined stress of the vertical water flow on the strain gage and the internal torque is:

$$\sigma_v = \frac{M_v x}{I} \tag{4}$$

Where x is the distance from the neutral axis and I is the moment of inertia of the cross section about the neutral axis.

Combining(2)(3)(4), it follows that

$$\sigma_v = \frac{1}{4I} C_D \rho v_v^2 b_1 l_1^2 h_1 \tag{5}$$

So:

$$\sigma_v \propto v_v^2 \tag{6}$$

Second, in the horizontal direction, as shown in Fig. 2(c), the headstream plane is subject to the combined effect of positive tension F_h and bending tension M_h. Considering the boundary layer theory, the flow rate of water in the boundary layer $(0, \delta)$ is different from that in the normal flow zone (δ, l_2). Since the experimental environment is less disturbing than the real environment, the water motion near the sensor can be simplified to a flat plate laminar boundary layer flow with Reynolds number $Re_l < 5 \times 10^5$。At this time, the flow speed

$$v_y = \begin{cases} v_\infty \left(\frac{2y}{\delta} - \frac{y^2}{\delta^2} \right) & 0 < y < \delta \\ v_h & \delta < y < l_2 \end{cases} \tag{7}$$

where v_∞ is the velocity at infinity, i.e., v_h.

Combined with (2)(3)(7), the combined force F_h and moment M_h of the water flow in the vertical direction on the strain gauge are

$$\begin{aligned} F_h &= F_{h1} + F_{h2} \\ M_h &= M_{h1} + M_{h2} \end{aligned} \tag{8}$$

Where

$$\begin{aligned} F_{h1} &= \int_0^\delta \frac{1}{2} C_D \rho b_2 v_h^2 \left(\frac{2y}{\delta} - \frac{y^2}{\delta^2} \right)^2 dy \\ F_{h2} &= \frac{1}{2} C_D \rho b_2 (l_2 - \delta) v_h^2 \\ M_{h1} &= \int_0^\delta \frac{1}{2} C_D \rho b_2 v_h^2 \left(\frac{2y}{\delta} - \frac{y^2}{\delta^2} \right)^2 y dy \\ M_{h2} &= \frac{1}{4} C_D \rho b_2 (l_2^2 - \delta^2) v_h^2 \end{aligned} \tag{9}$$

In addition, the relationship between the combined stress σ_h, the internal moment M_h and combined force F_h of the water flow in the horizontal direction on the strain gauge is

$$\sigma_h = \frac{F_h}{S_1} + \frac{M_h h_1}{2I} \tag{10}$$

Combined with (8) and (10), it can be known that the relationship between σ_h and horizontal velocity v_h is:

$$\sigma_h \propto {v_h}^2 \tag{11}$$

Third, considering the influence of horizontal shear stress τ:

$$\tau = \frac{2\mu v_\infty}{\sigma} = \frac{2\mu v_h}{\sigma} \tag{12}$$

The total combined stress σ on the strain gauge is:

$$\sigma = \sigma_v + \sigma_h - \tau \tag{13}$$

Combining (1)(12)(13), we obtain：

$$\sigma \propto v^2 \tag{14}$$

And because the relationship between stress and strain is

$$\varepsilon = \frac{\sigma}{E} \tag{15}$$

where E is the modulus of elasticity of the material.

Combining (14)(15), we can obtain the quadratic relationship between strain gauge strain ε and water flow velocity v as

$$\varepsilon \propto v^2 \tag{16}$$

3 Data Calibration Algorithm

3.1 Sensor Data Acquisition System

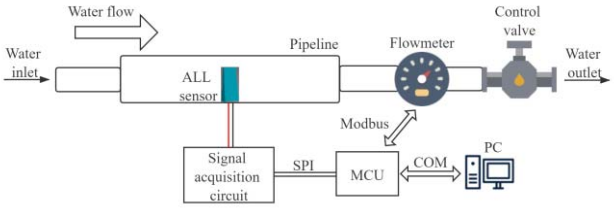


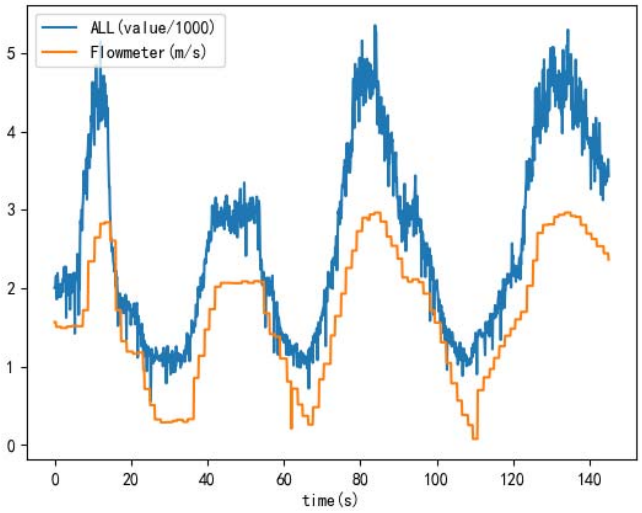
Fig 3: Sensor Data Acquisition System

To obtain the sensor data, we designed a data acquisition system as shown in Fig. 3. The ALL sensor was fixed in a horizontal pipe, and the water flow directly impacted the sensor under the drive of the pump to simulate the motion of the water relative to the robot fish. It should be noted that we did not stratify the water flow to retain the turbulent nature of the water flow, which helps us to simulate a realistic underwater turbulent environment[11]. The electrical signal from the sensor output is passed to the MCU through a signal acquisition circuit. We installed a commercial flow meter (ECLWGY50BLC1SSNM3) to be used as a calibration instrument. When the measured liquid flows through the flowmeter, the impeller is rotated by force and its rotational speed is proportional to the average flow rate of the pipe, which is then converted into an electrical signal by the magnetic field. The flow meter communicates with the MCU via Modbus. the MCU transmits the measured values of the ALL sensor and the flow meter to the PC via COM for subsequent data analysis. The flow rate of the water is manually controlled by a speed control valve. In this experiment, the sampling rate of both the ALL and the flowmeter is set to 11.8Hz.

3.2 Sensor Description of the raw sensor data

The data measured by the system mentioned above are shown in Fig. 4.

It can be clearly seen that there is significant noise in the ALL raw data due to the presence of turbulence, mechanical vibration, and other noise effects in the pipe. And the current noise model is very difficult to model accurately, how to achieve the calibration from sensor signal to flowmeter data (i.e., real velocity) without filters is a major achievement of this paper.



3.3 Data Calibration Algorithm

For a stable experimental system, the information such as the intrinsic noise in the ALL raw data measured by intercepting a small period of time should also be continuous. Analogous to the multidimensional time series prediction algorithm[12], this paper proposes a calibration mechanism based on sequence characteristics, i.e., ten successive measured ALL sensor data are viewed as a sequence to predict a flowmeter value (i.e., true velocity) at the same moment as the last data in that segment. Recurrent neural network RNN is mostly used for sequence learning because of its "memory", and its improved model LSTM is proven to effectively solve the "long environment" problem caused by RNN, because it saves important features of the sequence through various gate structures. So the LSTM model is chosen for the main network of the calibration algorithm in this paper. Since the weight of the historical sensor information is different for the current real velocity representation, the attention mechanism is introduced in the model. In addition, to effectively extract local features and prevent model overfitting, we add a convolutional layer to the model. The specific algorithm is as follows:

First, the model input data takes the following form.

$$T=\{(x_1,x_2,x_3,...,x_{10},y_{10}),$$
$$(x_2,x_3,x_4,...,x_6,y_{11}),$$
$$.....$$
$$(x_{N-9},x_{N-8},x_{N-7},...,x_N,y_N)\}$$

where x_i is the ALL sensor data, y_i is the flow meter data, and $i=1,2,3...,N$

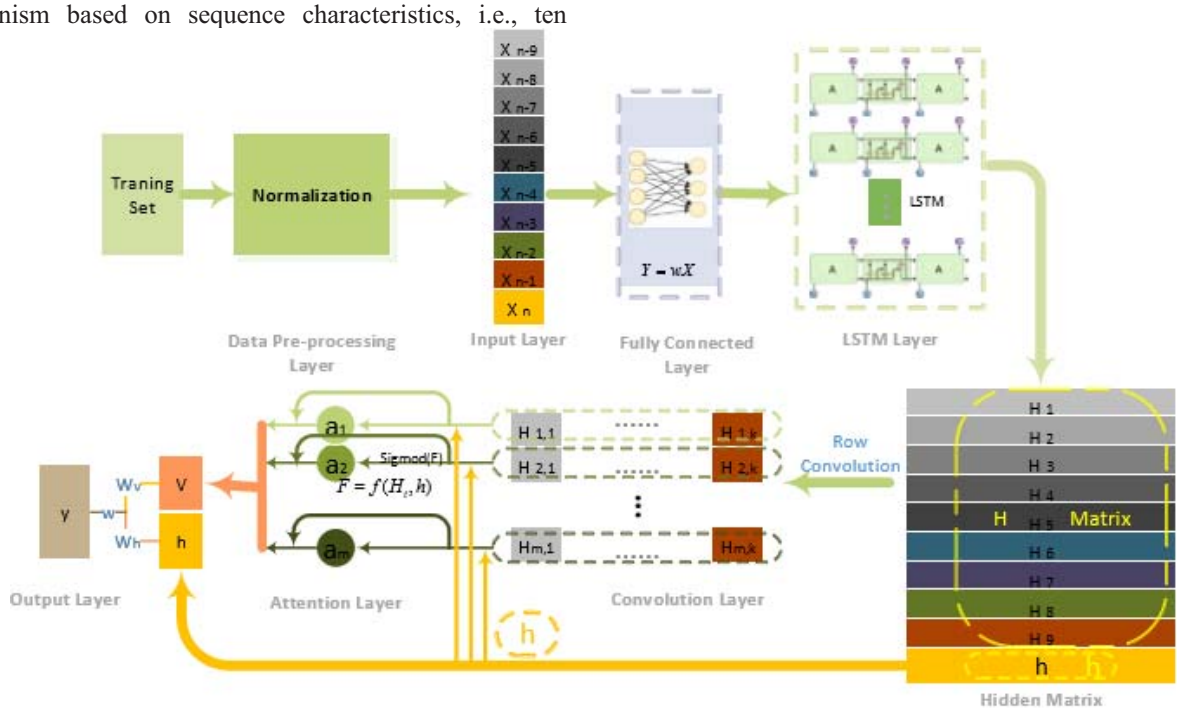


Fig. 6: Data calibration model of this paper

Second, the data calibration algorithm model proposed in this paper is shown in Fig. 6, which mainly consists of fully connected layers, LSTM layers, a convolution layer, and an Attention layer, where:

(1)Normalization formula (Z-Score normalization)

$$x_n=\frac{x-\mu}{\sigma+10^{-4}} \qquad (17)$$

Where, x is the original data, μ is the mean of x, and σ is the standard deviation of x.

(2)Convolution layer (row convolution of H-matrices)

$$H_i=\sum_{l=1}^{w}H_{i,w-l}\times C_l \qquad (18)$$

where the convolution kernel size is $(1,w)$ and $H_{i,j}$ denotes the resultant value of the action of the i-th row vector and the convolution kernel.

(3)Attention-weighting factor formula

$$a(i)=sigmoid((H_i)^T\cdot(W_a h)) \qquad (19)$$

where W_a is the learning parameter

(4)Context vector x

$$V=\sum_{i=1}^{m}a(i)H_i \qquad (20)$$

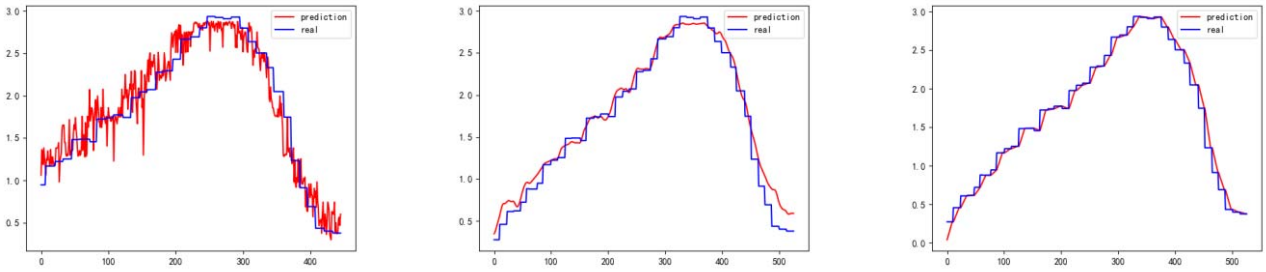
(5)Output layer (fusion of attention layer result and LSTM layer last layer result vector)

$$y=w(W_h h+W_V V) \qquad (21)$$

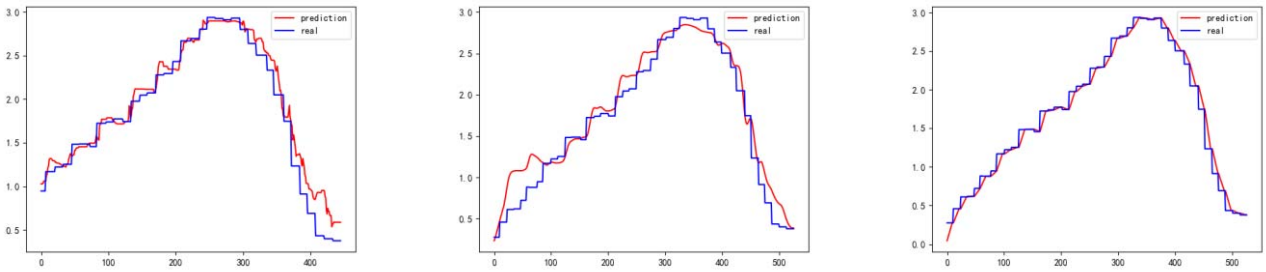
where w W_h W_V are learning parameters

3.4 Data Calibration Results

As shown in Fig. 4, 5, 6 and 7, in order to verify the effectiveness of the above algorithm, this paper controls the inlet valve of the sensor data acquisition system to change the flow velocity, making it approximately present the signal characteristics of sine, square, step and sine-square mixture wave. They are sequentially used as the signal input of the system, and four sets of ALL sensor raw data and corresponding flow meter data are collected. These data were used as calibration raw data and input to the model to test the calibration effect. It should be noted that because the inlet valve is difficult to control, the waveforms of these four signals are not very standard but this does not have any negative effect on the inspection results and is even more conducive to the robustness test. Meanwhile, in order to verify the superior robustness of the calibration model proposed in this paper, we improve the more popular regression algorithms such as Random Forest algorithm[13] and LSTM network[10] to do the calibration function on the same four sets of data to obtain comparative experimental results. In addition, another major advantage of the model in this paper is that there is no need to do filtering on the original data, and to verify this, comparative processing was also done within each group of experiments, separately with and without filter conditions. To ensure that the shape and width of the signal remain unchanged, the Savitzky-Golay filter (S-G filter) was used for the simple filter.

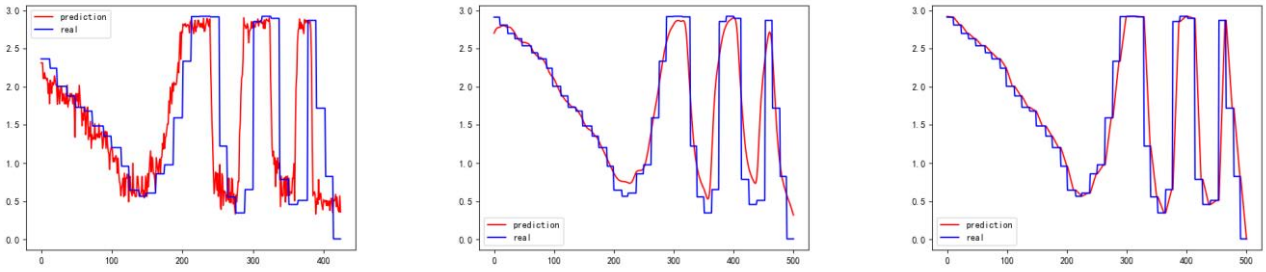


(a.1) Random Forest (without filter) (b.1)LSTM(without filter) (c.1) Algorithm of Us(without filter)

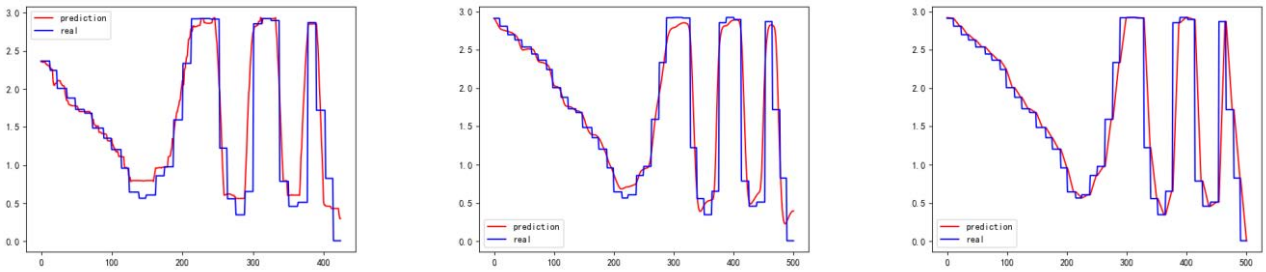


(a.2) Random Forest (S-G filter) (b.2)LSTM(S-G filter) (c.2) Algorithm of Us(S-G filter)

Fig. 4: Experimental results of approximate sinusoidal signal

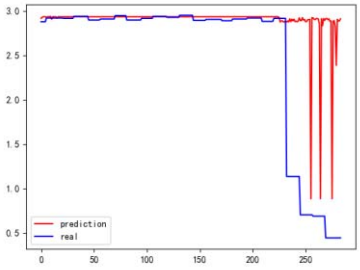


(a.1) Random Forest (without filter) (b.1)LSTM(without filter) (c.1) Algorithm of Us(without filter)

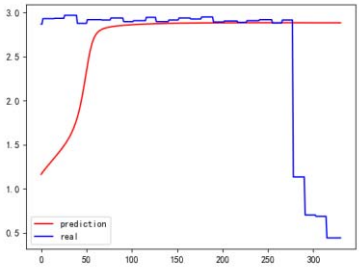


(a.2) Random Forest (S-G filter) (b.2)LSTM(S-G filter) (c.2) Algorithm of Us(S-G filter)

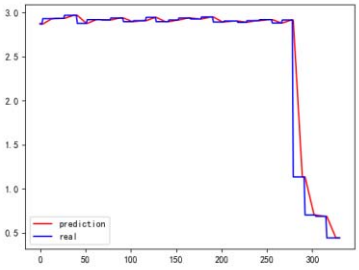
Fig. 5: Experimental results of approximate square wave signal



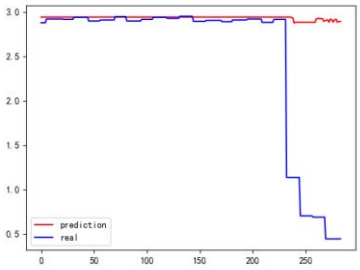
(a.1) Random Forest (without filter)



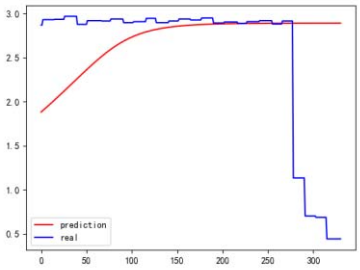
(b.1)LSTM(without filter)



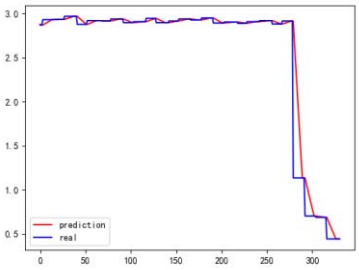
(c.1) Algorithm of Us(without filter)



(a.2) Random Forest (S-G filter)

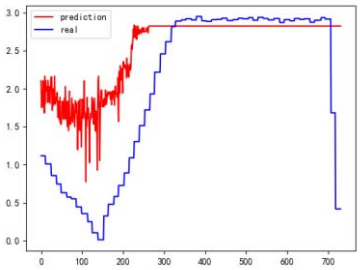


(b.2)LSTM(S-G filter)

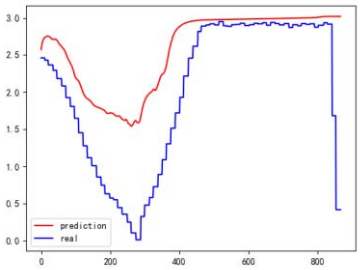


(c.2) Algorithm of Us(S-G filter)

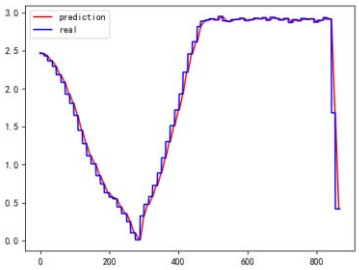
Fig. 6: Experimental results of approximate step signal



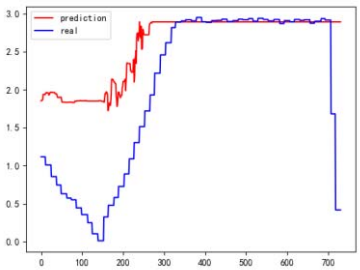
(a.1) Random Forest (without filter)



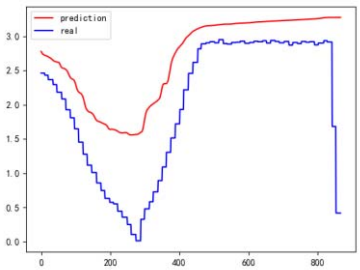
(b.1)LSTM(without filter)



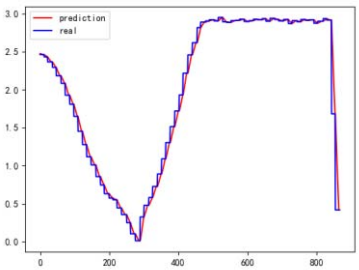
(c.1) Algorithm of Us(without filter)



(a.2) Random Forest (S-G filter)



(b.2)LSTM(S-G filter)



(c.2) Algorithm of Us(S-G filter)

Fig. 7: Experimental results of mixed signal

For our data calibration task, the constructed model belongs to a type of regression model, so its final model evaluation method can be used to evaluate linear regression models. In this experiment, MSE and R-square are chosen as our judging criteria.

In the evaluation process, the closer the MSE is to 0 and the closer the R-square is to 1, the better the regression of the model is, which also means the better the data calibration of our sensor is. The experimental results are shown in Tables 1, 2 and 3.

Table 1 Calibration Results of Random Forest Data Calibration Algorithm

Input Signal	No Filter		S-G Filter	
	MSE	R-Square	MSE	R-Square
Sine Signal	0.045342	0.923456	0.025603	0.971846
Square Signal	0.803615	0.240406	0.172845	0.801558
Step Signal	0.806957	0.097975	0.874318	0.752483
Mixture Signal	0.761276	0.751330	0.814484	0.824369

Table 2 Calibration Results of LSTM Data Calibration Algorithm

Input Signal	No Filter		S-G Filter	
	MSE	R-Square	MSE	R-Square
Sine Signal	0.018508	0.986404	0.035622	0.967714
Square Signal	0.138110	0.857803	0.109468	0.878251
Step Signal	1.064444	0.040884	0.888393	0.082065
Mixture Signal	0.668421	0.820668	0.714960	0.834069

Table 3 Calibration Results of Our Paper Data Calibration Algorithm

Input Signal	No Filter		S-G Filter	
	MSE	R-Square	MSE	R-Square
Sine Signal	0.009345	0.983824	0.009351	0.986813
Square Signal	0.176548	0.800098	0.176548	0.800098
Step Signal	0.039993	0.942434	0.039233	0.943505
Mixture Signal	0.017421	0.982770	0.017258	0.982921

From the experimental results, it is obvious that the presence or absence of a filter has a very significant impact on the random forest data calibration algorithm, and the calibration effect is basically not achieved without a filter. Although LSTM data calibration algorithm can achieve calibration without filter, its robustness cannot be guaranteed, and the calibration effect for step signal and mixed signal is too poor. However, the calibration algorithm proposed in this paper does not affect the calibration results with or without a filter. At the same time, for the four input signals, the calibration error is small and meets the calibration requirements.

4 Summary and future works

In order to solve the problem of swimming speed measurement of robot fish in complex underwater environment, this paper uses microelectronics technology to design a cantilever beam type ALL sensor monomer based on the piezoresistive effect by imitating the structure of Superificial Neuromasts (SNs) of real fish, and establishes a more accurate mechanism model for the sensor considering the complex effects of underwater turbulence and boundary layer. A more accurate mechanism model is established for this sensor considering the complex effects of underwater turbulence and boundary layer, and the sub-square relationship between the strain of the sensor core strain gauge and the water flow velocity is obtained, which is a very important guideline for the sensor material and selection. In addition, considering the complexity and chaos of physical modeling, we developed an "end-to-end" data calibration model based on the idea of sequence feature prediction, which can directly calibrate the original data without filters and shows superior performance in terms of robustness and accuracy.

At present, the ALL sensor proposed in this paper can already be well implemented for stationary machine fish fixed flow water velocimetry. For the real swimming machine fish, it is still necessary to form the sensor array, i.e., the artificial lateral line system, by the monomer in order to realize the accurate velocimetry for it. In the future research, we will build on the sensor and calibration model in this paper to realize the array data calibration of the artificial lateral line system to measure the velocity of the robot fish when it is swimming freely underwater.

References

[1] Freeman, W., *The function of the lateral line organs.* Science, 1928. **68**: p. 205-205.

[2] Mogdans, J. and H. Bleckmann, *Coping with flow: behavior, neurophysiology and modeling of the fish lateral line system.* Biological Cybernetics, 2012. **106**(11-12): p. 627-642.

[3] Asadnia, M., et al., *Flexible and Surface-Mountable Piezoelectric Sensor Arrays for Underwater Sensing in Marine Vehicles.* Ieee Sensors Journal, 2013. **13**(10): p. 3918-3925.

[4] Dagamseh, A., et al., *Imaging dipole flow sources using an artificial lateral-line system made of biomimetic hair flow sensors.* Journal of the Royal Society Interface, 2013. **10**(83).

[5] Dagamseh, A.M.K., et al., *Towards a high-resolution flow camera using artificial hair sensor arrays for flow pattern observations.* Bioinspiration & Biomimetics, 2012. **7**(4).

[6] Liu, P., R. Zhu, and R.Y. Que, *A Flexible Flow Sensor System and Its Characteristics for Fluid Mechanics Measurements.* Sensors, 2009. **9**(12): p. 9533-9543.

[7] Wang, C.C., W. Wang, and G.M. Xie, *Speed Estimation for Robotic Fish Using Onboard Artificial Lateral Line and Inertial Measurement Unit.* 2015 Ieee International Conference on Robotics and Biomimetics (Robio), 2015: p. 285-290.

[8] Xu, D., et al., *Hydrodynamic Analysis with an Artificial Lateral Line of Robotic Fish.* Proceedings of the 2017 12th Ieee Conference on Industrial Electronics and Applications (Iciea), 2017: p. 1380-1385.

[9] Wang, W., et al., *Speed Evaluation of a Freely Swimming Robotic Fish with an Artificial Lateral Line.* 2016 Ieee International Conference on Robotics and Automation (Icra), 2016: p. 4737-4742.

[10] Fuentes-Perez, J.F., et al., *Underwater vehicle speedometry using differential pressure sensors: Preliminary results.* 2016 Ieee/Oes Autonomous Underwater Vehicles (Auv), 2016: p. 156-160.

[11] Shih, S.Y., F.K. Sun, and H.Y. Lee, *Temporal pattern attention for multivariate time series forecasting.* Machine Learning, 2019. **108**(8-9): p. 1421-1441.

[12] Svetnik, V., et al., *Random forest: a classification and regression tool for compound classification and QSAR modeling.* J Chem Inf Comput Sci, 2003. **43**(6): p. 1947-58

[13] Hochreiter, S. and J. Schmidhuber, *Long short-term memory.* Neural Comput, 1997. **9**(8): p. 1735-80