ICIC Express Letters
Part B: Applications
Volume **11**, Number **3**, March **2020**

pp. **245–251**

DETECTION OF SWIMMING STROKE START TIMING BY DEEP LEARNING FROM AN INERTIAL SENSOR

Yuto Omae[1], Masahiro Kobayashi[2], Kazuki Sakai[3], Takuma Akiduki[4]
Akira Shionoya[2] and Hirotaka Takahashi[2]

[1]Department of Industrial Engineering and Management
College of Industrial Technology
Nihon University
1-2-1, Izumi, Narashino, Chiba 275-8575, Japan
yuto.omae@gmail.com

[2]Department of Information and Management System Engineering
Nagaoka University of Technology
1603-1, Kamitomioka, Nagaoka, Niigata 940-2188, Japan
hirotaka@kjs.nagaokaut.ac.jp

[3]Department of Electronic Control Engineering
National Institute of Technology, Nagaoka College
888, Nishikatakai, Nagaoka, Niigata 940-8532, Japan

[4]Department of Mechanical Engineering
Toyohashi University of Technology
1-1, Hibarigaoka, Tenpakucho, Toyohashi, Aichi 441-8580, Japan

Received September 2019; accepted December 2019

Abstract. *We aim for developing a system for supporting swimmer with an inertial sensor. It is required to develop a method to detect swimming stroke starting timing from the single inertial sensor data attached to swimmer's back waist (3-axis acceleration and gyro data). In this paper, we propose a detection method of swimming stroke starting timing from the time and frequency domain of 3-axis acceleration and gyro data by using the deep learning and peak detection. To learn the deep learning parameter, we carried out an experiment. The number of subjects is six and they swim butterfly. We assigned four persons as the training dataset and two person as the test dataset. As a result, our method achieves the high quality detection, i.e., the precision and recall are 0.855 and 0.904, respectively. Therefore, we confirm that there is a possibility that our method can use our system for supporting swimmer with an inertial sensor. Moreover, our method can be used to an estimation of starting timing of other human motions.*
Keywords: Inertial sensor, Human motion, Deep learning, Swimming motion

1. **Introduction.** Because it is important to promote swimming performance enhancement in Japan, we aim to develop a swimming performance analysis system [1].

In this system, single inertial sensor is attached to the swimmer's back waist. The system has a function that can understand respective stroke's performance from the sensor data (3-axis acceleration and gyro data). We have to develop an algorithm to detect the starting and ending timing of respective stroke to achieve it. Because the swimming stroke motion is cyclic, the ending timing of n is the starting timing of $n+1$ times stroke. Thus, we develop an algorithm of the start timing detection of respective stroke from the sensor data.

Many researchers performed to develop the swimming styles classification (back stroke, front stroke, breast stroke and butterfly) [2, 3, 4, 5, 6]. Moreover, Davey et al. developed an algorithm to count the number of strokes [7]. Jensen et al. and Kobayashi et al.

DOI: 10.24507/icicelb.11.03.245

developed a classification system that can detect the turn motion [4, 8]. Their researches have high quality. However, there are few researches of the swimming stroke by the sensor data. It is important to develop not only the counting strokes and swimming style classification system but also the accurate stroke performance evaluation method.

Recently, the deep learning technique is used to the research of various fields to achieve the high quality. Thus, many researchers introduce the human motions analysis based on the deep learning and the inertial sensor data. For example, Ordones and Daniel [9], Alsheikh et al. [10] and Ravi et al. [11] developed the human activities recognition system by using the convolutional neural network (CNN) and deep Boltzmann machine. Their researches' target is the basic human motion (e.g., walking, running). We focus on the application of the deep learning to the sports motion. In this paper, we focus on the swimming stroke and our target swimming style is butterfly, as the first step. Especially, in order to use our system for supporting swimmer with an inertial sensor, we propose a detection method of swimming stroke starting timing by using the deep learning and peak detection. There are researches of rough detection of swimming motions and styles by using inertial sensors and machine learning [3, 4, 5, 6, 8]. In contrast, our method aims for the detection of exact timing of swimming motion.

The paper is organized as follows. In Section 2, we give an overview of a detection method of swimming stroke starting timing which we use in this paper. In Section 3, we overview the experiment to verify the reliability of our detection method. Moreover, its result and discussion are presented. Section 4 is devoted to a summary.

2. **Outline of a Detection Method.** In this subsection, we briefly describe our proposed detection method of swimming stroke starting timing from the time and frequency domain of 3-axis acceleration and gyro data by using the deep learning and peak detection.

2.1. **Inertial sensor and its attached location.** We use a waterproofed inertial sensor shown in the right side of Figure 1: size: 67mm × 26mm × 8mm, weight: 20g, acceleration range: ±5G, gyro range: ±1500dps and sampling frequency: 100Hz developed by SPORTS SENSING Co., LTD [1, 6, 12]. Attached location of the sensor is swimmer's back waist. We used a double sided tape to attach an inertial sensor. Setting of the sensor position and the axis with the positive direction of arrow direction of acceleration and angular velocity are shown in Figure 1. Here, X_{acc} denotes X-axis acceleration, X_{ang} denotes X-axis gyro element and Y and Z axes are the same. From the measured data of inertial sensor, our proposed method detects the starting timing of respective stroke. In the remaining of this section, we explain the algorithm of our detection method.

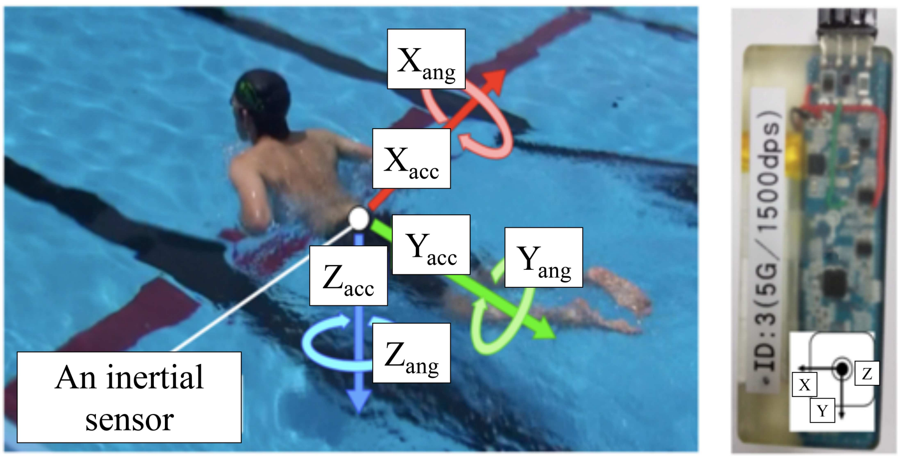

FIGURE 1. Attached location of single inertial sensor

2.2. **Construction of a CNN model.** Our algorithm extracts sub data from the sensor data. The left side of Figure 2 shows an example of our input data into the convolutional neural network (CNN). The line size is the "window length". Its length is the half value of average time of the strokes of butterfly. Because we know the average time m is 1.06sec [6], we set 0.53sec as its window length. Because the sampling frequency is 100Hz, 0.53sec means 53 size sequence data. The row size is twelve: the time and frequency domain data of 6 axis (X, Y, Z of acceleration and gyro data). The frequency domain is calculated by fast Fourier transform (FFT). As the result, 53×12 size data is inputted into CNN. Note that the 53rd line means the point of the starting timing. Therefore, the CNN detects the starting timing by using the sensor data of second half of stroke.

We describe the construction of a CNN model. Our model consists of two convolutional layers and three fully connected layers shown in the middle of Figure 2. The first and second convolutional layer are 3×1 filter and the ReLu function. The inputted data's size are 51×12 and 49×12 by the first and second convolutional layers, respectively. The neuron size of the first fully connected layer is 588 (49×12) and the activation function is the ReLu function. The neuron size of the second fully connected layer is 10 and the activation function is the ReLu function. The neuron size of the final fully connected layer is 3 and the activation function is the Softmax function. This is detection layers: the first neuron, second and third neuron mean "starting timing of stroke" (class c_1), "stroke (not starting timing)" (class c_2) and "turn or not swimming" class (class c_3), respectively (the right side of Figure 2). If the first neuron's value is the maximum, it means this motion is the starting timing of the stroke. If the second neuron's value is the maximum, it means this motion is the stroke that is not starting timing. If the third neuron's value is the maximum, it means this motion is not stroke (turn, walking, jumping and so on).

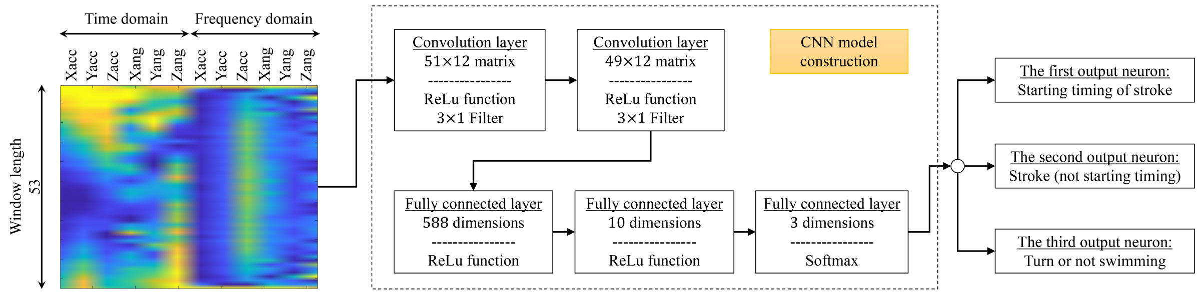

FIGURE 2. CNN model to detect the starting time of respective swimming stroke from single inertial sensor data

The above procedures are performed every 0.01sec by sliding window method [13, 14].

2.3. **Peak detection.** There is a possibility of the stroke start timing, if the output value of the first neuron is maximum. The sensor data is inputted into the CNN every 0.01sec because the sampling frequency is 100Hz. Therefore, the output values of the first neuron are formed of the time sequence data. To detect the starting timing of respective stroke, we adopt a peak detection to the output values of the first neuron. We defined that the peak timings are the respective stroke starting timing. The peak condition is over 0.33 of the output value of the first neuron because the number of output neuron is three and the summation value of the all output neurons is one (the mean value of each output neuron is 0.33).

According to Omae et al. [6], the standard deviation of the butterfly's stroke time s is 0.04sec. Thus, we consider that $n+1$ times starting timing of stroke does not occur prior to $m - 3s$ from n times stroke starting timing ($m - 3s = 1.06 - 3 \times 0.04 = 0.94$sec).

By their conditions, we develop the algorithm to detect the starting timing of the butterfly stroke.

3. **Experiment.** To verify the reliability of our algorithm, we carried out an experiment. In this section, the outline of the experiment, the result and discussion are presented.

3.1. **Outline.** The subjects are six university students who belong to swimming clubs in university and have the experience of the butterfly. Single water proof inertial sensor is attached to their back waist. The length of the pool we use is 25m and the subjects swim one lap. In other words, their swimming distance is 50m. We record the scene of the experiment by video camera (30fps) and add the class labels c_1, c_2 and c_3, respectively. Thus, the dataset consists of the pair of the inertial sensor data and the class label. We search for the video frames of human motions defined as the class labels in all video frames for all subjects. Then, we give the class labels to time sequence inertial sensor data.

Next, we introduce the CNN's learning condition. The number of learning is 2×10^4 times, the optimization algorithm is the stochastic gradient descent, learning rate is 0.01. We split the dataset consisting of the sensor data of six subjects collected by the experiment for the parameter learning and model validation. The four subjects are assigned as the training dataset and the two subjects are assigned as the test dataset. The training dataset is used for the CNN parameter learning and the test dataset is used for the evaluation of our method.

3.2. **Result and discussion.** A result of the output by our detection method of swimming stroke starting timing is shown in Figure 3. The horizontal axis means time sequence and the vertical axis means the output values of the first neuron. The output values of the first neuron becomes high value when the timing is respective starting stroke. The

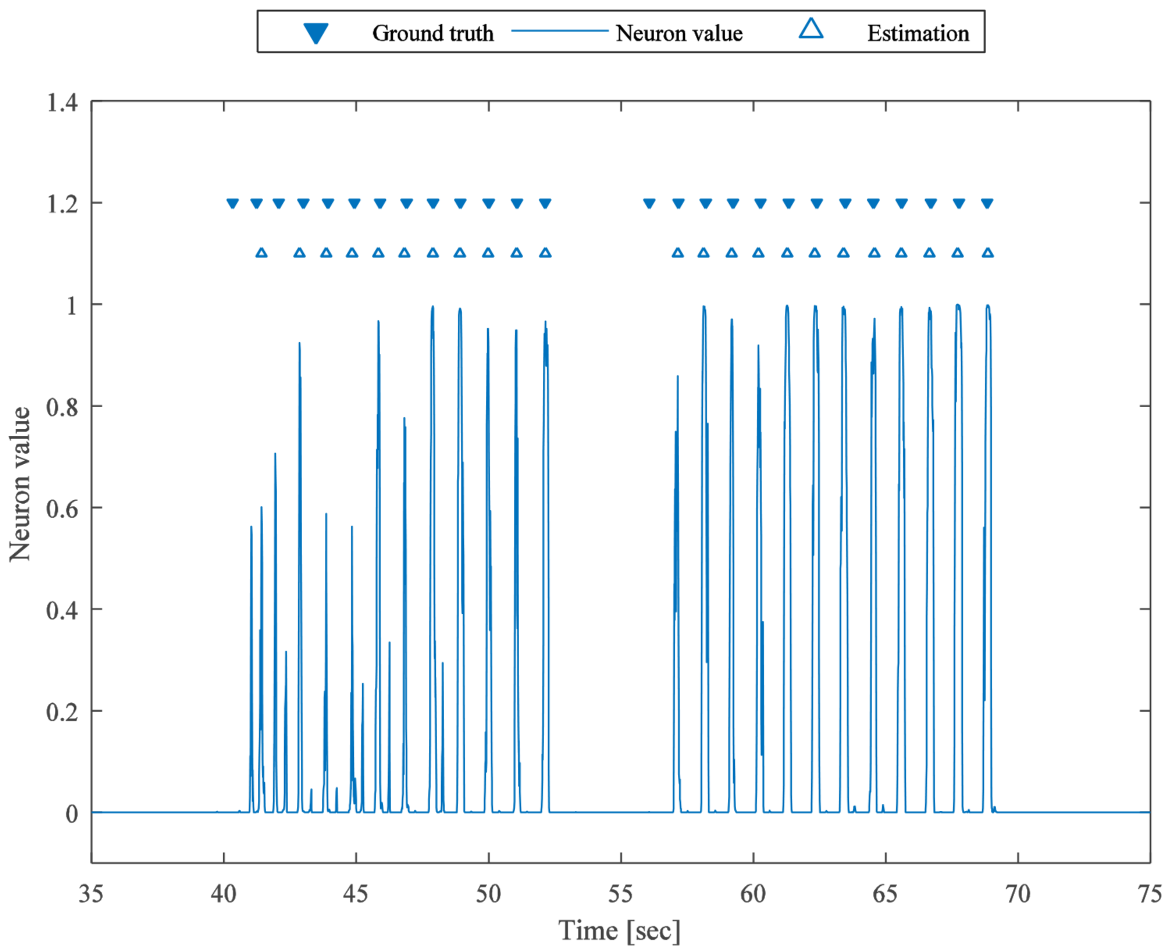

FIGURE 3. Time sequence of the output values of the first neuron, detection result and ground truth

filled triangles mean the peak detection results, i.e., the estimated stroke starting timings. The non-filled triangles mean the ground truth of the real stroke starting timings.

Until ~ 40 seconds, it means the time before swimming (walking, jumping, motion of going into the pool and so on). Therefore, the neuron values are almost zero. From 40 to 53 seconds, it means the first half of swimming (25m). Therefore, the neuron values have some responses. From 53 to 57 seconds, it means the time during the turn motion. Therefore, the neuron values are almost zero. From 57 to 69 seconds, it means the second half of swimming (25m). Therefore, the neuron values also have some responses. After 69 seconds, it means the time after swimming. Therefore, the neuron values are almost zero. From the filled and non-filled triangles, we can confirm that our method can detect stroke starting timing.

Next, we carried out the reliability survey of our method. This result is shown in Table 1. The result is verified by the test dataset that is not used to the parameter learning and consists of two subjects. a_1 means the number of correct estimation timing, a_2 means the number of not correct estimation timing, respectively. Moreover, b_1 is the number of succeeded detections of real timing and b_2 is the umber of failed detections of real timing, respectively. Specifically, if the differences both real timing and estimated timing is within ± 0.30sec, we considered that it is a correct answer.

Table 1. Evaluation of the reliability by using test dataset

Item	Value
a_1: Number of correct estimation timing	47
a_2: Number of not correct estimation timing	8
b_1: Number of succeeded detection of real timing	47
b_2: Number of failed detection of real timing	5
P: Precision:	0.855
R: Recall:	0.904

We also calculate the precision P and recall R. These are defined by:

$$P = a_1/(a_1 + a_2), \tag{1}$$

$$R = b_1/(b_1 + b_2). \tag{2}$$

The range of them is from zero to one and the high value means good method. From Table 1, the precision achieves 0.855 and recall achieves 0.904. There are a few miss detections. However, it shows our detection quality achieves good state. However, it is important to improve the reliability for users.

Moreover, we carry out the error evaluation survey to the strokes succeeded detection. This result is shown in Table 2. The mean value of the error achieves 0.169sec and its standard deviation achieves 0.037sec. It means our method achieved the high quality.

From these results, we confirm that there is a possibility that our method can use our system for supporting swimmer with an inertial sensor.

Table 2. Evaluation of the error by using the test dataset

Item	Value [sec]
Absolute mean of the error	0.169
Std value of the error	0.037
Minimum value of the error	0.050
Maximum value of the error	0.260

4. **Summary.** In this paper, we described a method to detect stroke starting timing by deep learning and peak detection. First, we extracted the acceleration and gyro signals by the sliding window method [13]. After that, the time and frequency domain by FFT were calculated. These data were inputted into the CNN every 0.01sec. We got the time sequence output of the first neuron for the stroke starting timing class c_1 of CNN and applied the peak detection to this output. As the result of the reliability survey by using the test dataset consisting of two subjects, there were a few miss detections. However, our method achieved the high quality detection, i.e., the precision and recall achieve 0.855 and 0.904, respectively. Thus, we confirmed that there was a possibility that our method could use our system for supporting swimmer with an inertial sensor.

We describe the future works: the number of subjects of the experiment that we carried out is six and very small. Therefore, we aim to increase the subjects to expand the training and test datasets to improve the detection quality. And, we try other swimming styles except butterfly (such as front crawl, backstroke and breast stroke). Moreover, we implement our algorithm to the swimming coaching system. In addition, we defined the correct detection as ± 0.30sec, subjectively. However, it is desirable to define the value by collecting opinions from the swimmers and coaches. This task is also one of the future works. Finally, we aim to apply our developed system to the swimming training environment.

Acknowledgment. This work was supported in part by JSPS Grant-in-Aid for Young Scientists (B) (Grant No. 17K13179 Y. Omae) and Young Scientists (Grant No. 19K20062; Y. Omae, Grant No. 19K14717; K. Sakai and Grant No. 19K14924; T. Akiduki). This work was also supported in part by JSPS Grant-in-Aid for Scientific Research (C) (Grant No. 16K06156; T. Akiduki and Grant No. 17K05437; H. Takahashi).

REFERENCES

[1] Y. Omae, Y. Kon, K. Sakai, H. Takahashi, T. Akiduki, C. Miyaji, Y. Sakurai, N. Ezaki and K. Nakai, Supporting environment to coach swimmer by data-driven approach, *JSME Symposium: Sports Engineering and Human Dynamics 2015*, A-14, 2015.

[2] Y. Ohgi, K. Kaneda and A. Takakura, Sensor data mining on the kinematical characteristics of the competitive swimming, *Procedia Engineering*, vol.72, pp.829-834, 2014.

[3] W. Choi, J. Oh, T. Park, S. Kang, M. Moon, U. Lee, I. Hwang and J. Song, Mobydick: An interactive multi swimmer exergame, *The 12th ACM Conference on Embedded Network Sensor Systems*, pp.76-90, 2014.

[4] U. Jensen, F. Prade and B. M. Eskofier, Classification of kinematic swimming data with emphasis on resource consumption, *2013 IEEE International Conference on Body Sensor Networks*, pp.1-5, 2013.

[5] P. Siirtola, P. Laurinen, J. Roning and H. Kinnunen, Efficient accelerometer-based swimming exercise tracking, *2011 IEEE Symposium on Computational Intelligence and Data Mining*, pp.156-161, 2011.

[6] Y. Omae, Y. Kon, M. Kobayashi, K. Sakai, A. Shionoya, H. Takahashi, T. Akiduki, K. Nakai, N. Ezaki, Y. Sakurai and C. Miyaji, Swimming style classification based on ensemble learning and adaptive feature value by using inertial measurement unit, *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol.21, no.4, pp.616-631, 2017.

[7] N. Davey, A. Megan and A. J. Daniel, Validation trial of an accelerometer-based sensor platform for swimming, *Sports Technology*, vol.1, pp.202-207, 2008.

[8] M. Kobayashi, Y. Omae, K. Sakai, A. Shionoya, H. Takahashi, T. Akiduki, K. Nakai, N. Ezaki, Y. Sakurai and C. Miyaji, Swimming motion classification for coaching system by using a sensor device, *ICIC Express Letters, Part B: Applications*, vol.9, no.3, pp.209-217, 2018.

[9] F. Ordones and R. Daniel, Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition, *Sensors*, vol.16, no.1, DOI: 10.3390/s16010115, 2016.

[10] M. A. Alsheikh, A. Selim, D. Niyato, L. Doyle, S. Lin and H. P. Tan, Deep activity recognition models with triaxial accelerometers, *The Workshops of the 30th AAAI Conference on Artificial Intelligence Applied to Assistive Technologies and Smart Environments*, Technical Report WS-16-01, pp.8-13, 2016.

[11] D. Ravi, C. Wong, B. Lo and G. Z. Yang, A deep learning approach to on-node sensor data analytics for mobile or wearable devices, *IEEE Journal of Biomedical and Health Informatics*, vol.21, no.1, pp.56-64, 2017.

[12] Sports Sensing Co., LTD., *9 Axis Waterproof Type Wireless Motion Sensor (Model Number: SS-WS1215, Type A)*.

[13] L. Bao and S. S. Intille, Activity recognition from user-annotated acceleration data, *Pervasive Computing*, pp.1-17, 2017.

[14] N. Ravi, N. Dandekar, P. Mysore and M. Littman, Activity recognition from acceleration data, *AAAI*, vol.5, pp.1541-1546, 2005.