https://doi.org/10.20965/jaciii.2017.p0616

Paper:

Swimming Style Classification Based on Ensemble Learning and Adaptive Feature Value by Using Inertial Measurement Unit

Yuto Omae*[1], Yoshihisa Kon*[2], Masahiro Kobayashi*[2], Kazuki Sakai*[3], Akira Shionoya*[2], Hirotaka Takahashi*[2], Takuma Akiduki*[4], Kazufumi Nakai*[5], Nobuo Ezaki*[5], Yoshihisa Sakurai*[6], and Chikara Miyaji*[7]

*[1]Department of Electrical Engineering, National Institute of Technology, Tokyo College
1220-2 Kunugida, Hachioji, Tokyo 193-0942, Japan
E-mail: yuto.omae@gmail.com
*[2]Department of Information and Management Systems Engineering, Nagaoka University of Technology
1603-1 Kamitomioka, Nagaoka, Niigata 940-2188, Japan
E-mail: hirotaka@kjs.nagaokaut.ac.jp
*[3]Department of Information Science and Control Engineering, Nagaoka University of Technology
1603-1 Kamitomioka, Nagaoka, Niigata 940-2188, Japan
*[4]Toyohashi University of Technology
1-1 Hibarigaoka, Tenpakucho, Toyohashi, Aichi 441-8580, Japan
*[5]National Institute of Technology, Toba College
1-1 Ikegamicho, Toba, Mie 517-8501, Japan
*[6]Sports Sensing Co., LTD.
2-2-1-403 Mukaino, Minami, Fukuoka 815-0035, Japan
*[7]Department of Creative Informatics, The University of Tokyo
7-3-1 Hongo, Bunkyo, Tokyo 113-8656, Japan

[Received October 22, 2016; accepted March 16, 2017]

We have been constructing a swimming ability improvement support system. One of the issues to be addressed is the automatic classification of swimming styles (backstroke, breaststroke, butterfly, and front crawl). The mainstream swimming style classification technique of conventional researches is based on non-ensemble learning; in their classification, breaststroke and butterfly are mixed up with each other. To improve its generalization performance, we need to use better classifiers and more adaptive feature values than previously considered. Therefore, this research has introduced (1) random forest technique, one of ensemble learning techniques, and (2) feature values specific to breaststroke and butterfly to construct a four-swimming-style classifier that has resolved this issue. From subjects with 7 to 20 years history of swimming races, we have obtained their sensor data during swimming and have divided the data into learning data and test data. We have also converted them into feature values that represent their body motions. We have selected from those body-motion-representing feature values the important data to classify four swimming styles and feature values specific to breaststroke and butterfly. We have used the learning data to construct a swimming style classifier, and the test data to evaluate its classification accuracy. The evaluation results show that (1') the introduction of ensemble learning has improved the mean value of F-measure for breaststroke and butterfly by 0.053, and (2') the introduction of feature values specific to breaststroke and butterfly has improved the mean value of F-measure for breaststroke and butterfly by 0.121 as compared with (1'). The proposed swimming style classifier has performed a mean F-measure of 0.981 for the four swimming styles as well as good classification accuracies for front crawl and backstroke. Therefore, we have concluded that the swimming style classifier we have constructed has resolved the problem of mixing up breaststroke and butterfly, as well as can properly classify all different swimming styles.

Keywords: swimming style classification, machine learning, ensemble learning, random forest, inertial measurement unit

1. Introduction

Swimming is designated by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) as "an event in which we can expect to win plural medals including gold medal (target event A)" [1]. To improve international swimming competition abilities, it is crucial to elevate the beginners and intermediate-level swimmers up to the top level as well as to develop top-level swimmers.

To improve swimming competition abilities, it is desirable to evaluate the swimming motion performance in minute detail. Generally, swimming motion performance is visually evaluated by a coach. However, the coach cannot evaluate the minute swimming motion in each stroke. These difficulties could be addressed to some extent by using a large-scale installation, such as motion capture, which is not applicable in general practical environment except to some limited number of swimmers at the top level.

In this context, we propose a swimming ability improvement support system using a compact sensor (hereinafter, a sensor) that can measure 3-axial acceleration, angular velocity and terrestrial magnetism (acceleration and angular velocity are hereinafter collectively called sensor data), can be used underwater, and can radio transmit data to a computer [2]. It would be more desirable to use many sensors to derive minute performance evaluation results, but the proposed system uses only one sensor so that the sensor installation should not give uncomfortable feeling to swimmers and that it could be easily used.

The proposed system aims to derive a swimmer's speed per stroke immediately after each stroke and then feed it back to the swimmer in the race via a receiver-installed bone conduction earphone or LED on a pair of goggles. It needs to be able to detect in minute detail the parts of the sensor data waveforms where the swimmer strokes. As stroking motions largely differ with swimming styles, sensor data waveforms when swimmers stroke are also different. To search with minimal errors the parts of sensor data waveforms that represent swimmers in stroking motion, it is desirable to use a one-stroke detection algorithm specific to a certain swimming style. This one-stroke detection algorithm can be used only if the swimming style is known. A swimmer or a coach could manually designate a swimming style. However, swimmers would find it burdensome to input a swimming style into a computer during each attempt in an environment where swimmers practice training without a coach or where they practice different swimming styles in a day. Therefore, it is desirable to be able to classify automatically the swimming styles.

The foregoing necessitates the construction of a swimming style classifier that can automatically classify swimming styles (backstroke, breaststroke, butterfly, and front crawl) from sensor data. The proposed system is aimed at real time feedbacks during a swimming race and needs to be able to identify swimming styles in the shortest time possible after the commencement of a swimming race rather than after the end of the race.

Many researches [3–6] are available discussing the classifications of these four swimming styles (see Section 2). However, they have the same problem in common where they mix-up breaststroke and butterfly. Therefore, to construct a practical swimming style classifier, we need to reduce the classification errors between breaststroke and butterfly.

To improve the classifier's generalization ability, it is essential to select better models and feature values than previously considered. Previous researches [3–6] all adopt non-ensemble learning techniques and often low-bias/high-variance models. The introduction of ensemble learning in the construction of swimming style classifiers should be able to reduce the model's variance and its classification errors accordingly. The introduction of feature values specific to breaststroke and butterfly could also resolve the problem with the previous researches.

Therefore, in this paper, we introduce (1) ensemble learning and (2) feature values specific to breaststroke and butterfly in constructing a four-swimming-style classifier that can properly classify all swimming styles. We examine the actual contribution of (1) and (2) in reducing classification errors. We also evaluate through simulation experiments the value of time needed by the proposed system after the start of swimming strokes to classify its swimming style.

This paper consists of seven sections and is organized as follows. Section 2 presents overviews of researches related to the classification of swimming styles. Section 3 discusses the data acquisition experiments and the technique for converting the acquired data into feature values, as well as the results of dividing the acquired data into learning and test data.

Section 4 presents the results of selecting some important feature values from those specified in Section 3 for the classification of swimming styles. It also reports the results of selecting some important feature values to classify breaststroke and butterfly so that the common problem of mixing up breaststroke and butterfly can be resolved. One of the differences between this paper and the previous researches is that we have selected some important feature values for the classification of breaststroke and butterfly and have constructed a space for these selected feature values to resolve the common problem with the previous researches.

Section 5 refers to the results of constructing swimming-style classifiers based on non-ensemble learning and ensemble learning. One of the differences between this paper and the previous researches is that while the previous researches have mainly classified swimming styles on non-ensemble learning, we have adopted ensemble learning to obtain better classification conditions resulting from reduced variances. In this section, we have verified with test data not involved in learning the contribution of (1) ensemble learning and (2) feature values specific to breaststroke and butterfly to the improvements in the classifier's generalization ability. For the verification, we have constructed four kinds of swimming style classifiers, namely, none of (1) and (2) are introduced; (1) only is introduced; (2) only is introduced; both (1) and (2) are introduced.

Section 6 proposes a technique to improve further the classifier's accuracy by applying a swimming style classifier with the introduction of (1) and (2). The results of the verification of the proposed technique's accuracy are also presented.

2. Related Researches

We present an overview of the previous researches on the construction of swimming style classifiers with a single sensor and refer to the differences between this paper and the previous researches.

Siirtola et al. [7] have attempted to construct a swimming style classifier capable of classifying backstroke, breaststroke and front crawl on the three-axial acceleration data as obtained from swimmers swimming with a sensor worn on the wrist or on the back. However, they have not attempted to classify butterfly.

Kon et al. [3] have attempted to construct a swimming style classifier on the three-axial acceleration data as obtained from swimmers swimming with a sensor worn on the waist. Their research used three kinds of feature values: average, variance and frequency domain entropy. They have used the Decision Tree (DT) as a model in constructing their classifier.

Ohgi et al. [4] have attempted to construct a swimming style classifier on the three-axial acceleration data as obtained from swimmers swimming with a sensor worn on the breast. Their research used three kinds of feature values: average, variance and skewness. They have used DT and Multiple Neural Network (MNN) as models in constructing their classifier.

Choi et al. [5] have attempted to construct a swimming style classifier on the three-axial acceleration/angular velocity data as obtained from swimmers swimming with a sensor-installed smartphone worn on the arm. Their research used four kinds of feature values: average, variance, minimum value and maximum value. They have used three kinds of models: DT, Support Vector Machine (SVM) using linear/non-linear kernels and Naive Bayes (NB) in constructing their classifier.

Jensen et al. [6] have attempted to construct a swimming style classifier on the three-axial acceleration/angular velocity data as obtained from swimmers swimming with a sensor worn on the back of the head. Their research used the following feature values: average, variance, kurtosis, skewness, minimum value and maximum value. They have used Linear Regression Model (LRM) as a model in constructing their classifier.

Any of these researches has the problem of mixing up breaststroke and butterfly in classifying swimming styles, which is probably because breaststroke and butterfly resemble each other in their motions. In backstroke, the swimmer's back faces to the floor and the stroke is alternately made by one arm after the other, while in front crawl, the swimmer's back faces to the ceiling and the stroke is alternately made by one arm after the other. In breaststroke and butterfly, the swimmer's back faces to the ceiling and the stroke is simultaneously made by both arms. This shows that there exist distinct differences in motions in the combination of swimming styles other than that of breaststroke and butterfly, while the body motions in the combination of breaststroke and butterfly resemble each other, which could result in obtaining similar data from their inertial measurement units and mixing up breaststroke and butterfly in the previous researches.

This paper discusses how to resolve the problem with many of the previous researches of mixing up breaststroke and butterfly by introducing (1) ensemble learning and (2) feature values specific to the classifications of breaststroke and butterfly.

Some researches on the classification of swimming styles with video cameras [8] are also available. The present research aims to construct an environment where any developed classifiers can be used as easily as possible as described in Section 1. We aim to achieve classifications of swimming styles only by averages of compact inertial measurement units without using a video camera.

3. Data Acquisition Experiments

Learning data are required for constructing a swimming style classifier and test data to obtain its generalization ability. This section describes the experiments to acquire data necessary for the construction of a swimming style classifier as well as the processes for constructing learning and test data.

3.1. Experimental Overview

We conducted experiments with a total of 13 university student subjects (9 males and 4 females) who belong to swimming clubs in the university. The subjects were 19.9 ± 1.7 years of age, 168.8 ± 6.6 cm in height, and 63.4 ± 5.3 kg in weight and have swimming experience as long as 14.2 ± 3.9 years. We conducted the experiments using a swimming pool of 25 m course and swimming club members only who have consented to our prior explanation that "the acquired data shall be used only for research purposes."

We used a sensor made by Sports Sensing Co., Ltd. (former Logical Products Co., Ltd.) with the following specifications: accelerations (± 5 G); angular velocities (± 1500 dps); terrestrial magnetism sensors; sampling frequency 100 Hz; weight 20 g; size 67 mm$\times$26 mm$\times$8 mm; and acquired data stored in a built-in memory (32 MB). For detailed specifications of the sensor, please refer to the product catalog [9] (under the official name: Water-Proof 9-Axis Wireless Motion Sensor (5 G/1500 dps), Model Number: SS-WS1215, Type A.

Figure 1 shows the position where the sensor is fitted and the axial setting, where X_{acc} denotes the X-axis accelerations and X_{ang} denotes the X-axis angular velocities. The same applies to Y and Z.

The subjects are instructed before the experiments to select two of the four swimming styles that they are good at and to do a lap swim of the 25 m course (50 m in total) with full force in the selected swimming styles without jumping into the pool. We have taken movies of their swimming with a video camera (30 fps) to compare their swimming motions with the sensor data waveforms. The video camera we have used is a digital HD video camera recorder HDR-CX720V made by Sony [10].

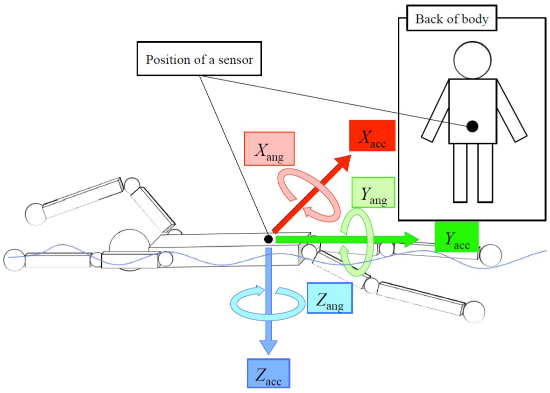

Fig. 1. Position to fit the sensor and axial setting.

Table 1. Average and standard deviation of stroke period (learning data).

Swimming style k	Number of strokes	m_k [s]	p_k
Backstroke (Ba)	32	1.254	0.121
Breaststroke (Br)	47	1.148	0.099
Butterfly (Bu)	67	1.060	0.040
Front crawl (Fr)	73	1.175	0.101

3.2. Data Processing

3.2.1. Division into Learning Data and Test Data

Of the 26 swimming trials (13 subjects×2 swimming styles), the data on five swimming trials are not available either because of a dead battery of the sensor or the sensor getting out of the lower back. Therefore, we have used the data on 21 swimming trials.

The data are broken down into three swims in backstroke, four swims in breaststroke, six swims in butterfly, and eight swims in front crawl. They are divided into data for the construction of a classifier (learning data) and data for the verification of its generalization ability (test data) by ensuring that the learning data and test data should not retain data on any similar subjects. As a result, the backstroke swimming learning data are for two swims and the test data are for one swim; the breaststroke swimming learning data are for two swims, and the test data are for two swims; the butterfly swimming learning data are for four swims, and the test data are for two swims; the front crawl swimming learning data are for five swims, and the test data are for three swims.

Furthermore, we have counted the number of the subjects' strokes on their movie data to calculate the total sum of strokes for each swimming style. As a result, the total number of backstrokes is counted as 52 (32 in learning data and 20 in test data); 85 for breaststrokes (47 in learning data and 38 in test data); 111 for butterfly (67 in learning data and 44 in test data); and 126 for front crawl (73 in learning data and 53 in test data).

3.2.2. Determination of Stroke Periods

In constructing a classifier to determine some motions on sensor data, we should not use raw data. Instead, we should convert them into variables that could represent the features of objects to be determined (hereinafter, feature values). In determining body motions on sensor data, their averages and/or variances in a certain length of time are used as feature values that represent average body swings or their distributions [11, 12]. The ideal length of time depends on the motions to be determined. Given that swimming race is composed of cyclic repetitions of stroke motions, we have decided to derive these values from their stroke periods.

Given that swimming in backstroke, breaststroke, butterfly and front crawl largely differs in stroke motions with each other, the widths of stroke periods must also be different from each other. Given that this research aims to construct a classifier capable of automatically classifying swimming styles, we should not use different stroke periods that suit different swimming styles. We need to derive a unique stroke period that could be partly applied to any kind of swimming styles. This subsection describes how to derive this unique stroke period.

At first, we have checked the movie data to determine the ordinary stroke starting point u on the sensor data. We have measured on the movie data the time each stroke has taken to calculate its mean value m_k [s] and standard deviation p_k. The results are shown in **Table 1**, where we derive a stroke period that would be partly applicable to any of the four swimming styles. Assuming that the stroke period r [s] of each swimming style follows a normalized distribution;

$$\mathcal{N}(r|m_k, p_k^2) = \frac{1}{\sqrt{2\pi p_k^2}} \exp\left(-\frac{(r-m_k)^2}{2p_k^2}\right). \quad (1)$$

With k denoting a swimming style, it is expressed as follows:

$$k \in C_{\text{Swim}} = \{\text{Ba}, \text{Br}, \text{Bu}, \text{Fr}\}, \quad (2)$$

where Ba denotes backstroke, Br, breaststroke, Bu, butterfly, and Fr, front crawl. An independent variable that yields the maximum value of Eq. (1) is considered a representative stroke period of the swimming style k. Given that this research aims to derive the width w of a stroke period that could be partly applied to any swimming styles, we calculate the total sum of the distributions in Eq. (1) to obtain the following equation:

$$f(r) = \alpha \sum_{k \in C_{\text{Swim}}} \mathcal{N}(r|m_k, p_k^2), \quad (3)$$

where α denotes a normalization constant where $f(r)$, integrated by its open interval $(-\infty, \infty)$, makes an integral value of 1. Given that we have summed four normal distributions, it becomes $1/4$.

As $f(r)$ is a function composed of the total sum of normal distributions of the stroke period of all swimming styles, we assume that its maximum value could be used as the width w of the stroke period that could be partly

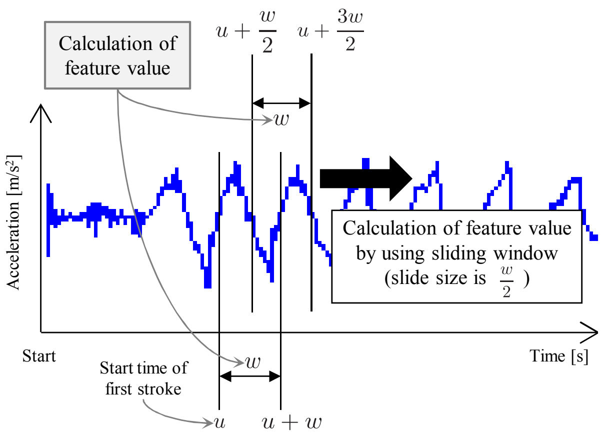

Fig. 2. Converting sensor data into feature values.

applied to any swimming style: w is defined as follows:

$$w = \underset{r}{\operatorname{argmax}} f(r). \quad (4)$$

Then, we obtained $w = 1.070$ s.

3.2.3. Conversion into Feature Values

This subsection describes the processes to convert sensor data into feature values using the width w of a stroke period that could be partly applied to any swimming style as obtained in the previous subsection.

Figure 2 shows the processes for converting sensor data into feature values. We have derived the feature values using the sliding window technique proposed by Bao et al. [11] and Ravi et al. [12].

The time-series data on the accelerations/angular velocities obtained for each swimming style and each subject are processed into feature values in the range from the ordinary stroke starting point u to $u+w$ in the first 25 m swimming. These processed feature values are shown in **Table 2**, where Mean denotes mean value; Var, variance; Skew, skewness; Kurt, kurtosis; Max, maximum value; Min, minimum value; Ent, frequency domain entropy. We have calculated these seven kinds of values for the three-axial accelerations and angular velocities and have adopted them as feature values $a_j (j = 1, \ldots, 42)$ by which to classify swimming styles. We have decided these feature values by referring to the previous researches [3–6] on the classification of swimming styles. As averaged angular velocities in any swimming styles tend to become 0, mean angular velocities cannot be used as feature values by which to classify swimming styles (for example, a swimmer tries to keep the body horizontal to the water surface by turning left after turning right). Therefore, we have converted angular velocities into absolute values before calculating their mean values. We have sequentially calculated the mean values in the process by sliding it by $w/2$ from the stroke's starting point u until immediately before turning back. We have applied the same processes to the strokes in the latter 25 m swimming. We have used only one graph (**Fig. 2**) to explain the processes,

Table 2. Adopted feature values $a_j (j = 1, \ldots, 42)$.

Definition	Acceleration j	Acceleration k	Angular Velocity j	Angular Velocity k
Mean value	1	X_{acc}	22	X_{ang}
$a_j = \text{Mean}(k)$	2	Y_{acc}	23	Y_{ang}
	3	Z_{acc}	24	Z_{ang}
Variance	4	X_{acc}	25	X_{ang}
$a_j = \text{Var}(k)$	5	Y_{acc}	26	Y_{ang}
	6	Z_{acc}	27	Z_{ang}
Skewness	7	X_{acc}	28	X_{ang}
$a_j = \text{Skew}(k)$	8	Y_{acc}	29	Y_{ang}
	9	Z_{acc}	30	Z_{ang}
Kurtosis	10	X_{acc}	31	X_{ang}
$a_j = \text{Kurt}(k)$	11	Y_{acc}	32	Y_{ang}
	12	Z_{acc}	33	Z_{ang}
Maximum value	13	X_{acc}	34	X_{ang}
$a_j = \text{Max}(k)$	14	Y_{acc}	35	Y_{ang}
	15	Z_{acc}	36	Z_{ang}
Minimum value	16	X_{acc}	37	X_{ang}
$a_j = \text{Min}(k)$	17	Y_{acc}	38	Y_{ang}
	18	Z_{acc}	39	Z_{ang}
Frequency domain	19	X_{acc}	40	X_{ang}
entropy $a_j = \text{Ent}(k)$	20	Y_{acc}	41	Y_{ang}
	21	Z_{acc}	42	Z_{ang}

M_{acc} : M-axial acceleration
M_{ang} : M-axial angular velocity

but in practice, there are six graphs in total representing the combinations of the $X/Y/Z$ axes and accelerations/angular velocities. We have applied the processes to the sensor data divided into learning data and test data to obtain learning data L and test data T that retain the converted feature values. Learning data L and test data T are labeled to represent swimming styles (backstroke, breaststroke, butterfly, and front crawl).

The processes using a sampling frequency of 100 Hz should make the second decimal place of w an even number. Therefore, with $w = 1.070$ s, we use 1.080 s.

The processes applied to the sensor data have resulted in the number of samples N_L of the learning data at 412 (breakdown: 60 in backstroke, 90 in breaststroke, 126 in butterfly, 136 in front crawl). The number of samples N_T of the test data is at 294 (breakdown: 38 in backstroke, 72 in breaststroke, 84 in butterfly, 100 in front crawl). In this research, we use the learning data L to construct a four-swimming-style classifier. We also use the test data T to verify its generalization ability for possible applications to unknown third parties.

4. Selection of Feature Values

The construction of the classifier with 42 types of feature values shown in **Table 2** will create a 42-dimensional vast feature-value space. This a high-dimensional vast feature-value space is not so desirable from the generalization point of view that it should be made lower dimensional by calculating the importance of feature values on some indices to select the feature values to be used rather than using all of the 42 kinds of feature values. It is important in the problem of multi-class classifications to select the feature values that can properly classify all classes. Therefore, we first select feature values that are important for the classifications of four swimming styles, and then we select feature values specific to breaststroke and butterfly swimming styles, which the previous researches have often misclassified.

4.1. Random Forest and Decision Tree

Among several available techniques of calculating the importance of feature values, we have adopted a technique that applies bootstrap sampling and random forest (RF). This subsection describes RF and decision tree (DT).

At first, we divide the learning data L generated in Subsection 3.2.3 by the bootstrap sampling as follows:

$$L^{\mathrm{sub}} = \{L_i^{\mathrm{sub}}\}_{i=1}^{B}, \quad (5)$$

$$O = \{O_i\}_{i=1}^{B}, \quad (6)$$

where L^{sub} denotes the sub-learning data with subsets of L as components; L_i^{sub} denotes learning data for constructing i-th weak learner; O denotes out of bag (OOB) dataset with subsets of L as components. O_i is composed of data that have not been selected as components of L_i^{sub}. OOB dataset O can be used for the parameter-tuning intended to improve the generalization performance of the weak learner composed of sub-learning dataset L^{sub} and for calculating the importance of feature values. B denotes the number of weak learner to be constructed. Next, we construct a weak learner using L^{sub} composed by the bootstrap sampling as follows:

$$Tr = \{Tr(i)\}_{i=1}^{B}, \quad (7)$$

where Tr denotes weak learner set. Its component $Tr(i)$ denotes weak learner constructed with i-th sub-learning data L_i^{sub}. In this research, we have decided to set the weak learner model as DT and the classifier as RF that decides its final classifications on the majority of DTs.

DT is a technique of representing classification standard by a binary tree structure and is constructed as follows. Ordinary DT is processed with learning data L and DT as weak learner of RF, with sub-learning data L_i^{sub}. First, the information entropy of samples in the undivided conditions is calculated:

$$I_{\mathrm{pre}} = -\sum_{k \in C_{\mathrm{Swim}}} p(k) \log p(k), \quad (8)$$

where $p(k)$ denotes the percentage of swimming style k. Then, divide the samples into two branches, namely, $a_j \geq s$, $a_j < s$, using a feature-value a_j. Nodes generated in the process are denoted by N_1 and N_2, and the numbers of samples by n_1 and n_2 samples. The information entropy with each of the divided nodes is calculated and averaged as follows:

$$I_{\mathrm{post}}^{a_j,s} = -\sum_{b=1}^{2} \frac{n_b}{n_1+n_2} \sum_{k \in C_{\mathrm{Swim}}} p(k|N_b) \log p(k|N_b). \quad (9)$$

This equation represents the fuzziness of each swimming style in each of the nodes divided by $a_j \geq s$, $a_j < s$. Then, the information gains are calculated as follows:

$$G(I_{\mathrm{pre}}, I_{\mathrm{post}}^{a_j,s}) = I_{\mathrm{pre}} - I_{\mathrm{post}}^{a_j,s}. \quad (10)$$

This equation serves as an index to measure how much fuzziness has been reduced before and after dividing the samples by $a_j \geq s$, $a_j < s$. Swimming styles are classified by seeking feature values a_j with high information gains and their branching criteria s. Feature values $a_j^{\max}$ and branching criteria $s^{\max}$, which can best classify swimming styles after being divided into two branches, can be obtained by solving the following optimization problem:

$$(a_j^{\max}, s^{\max}) = \underset{a_j,s}{\operatorname{argmax}}\, G(I_{\mathrm{pre}}, I_{\mathrm{post}}^{a_j,s}). \quad (11)$$

A binary tree is generated by recursively executing the processes. If a binary tree is made deeper under the said conditions, it will continue to be branched until one sample is generated, causing overfitting. Therefore, the following constraint conditions are imposed on the number of samples n_1 and n_2 after they are divided into two:

$$\text{subject to } d \leq n_1, n_2. \quad (12)$$

Setting these constraint conditions enable the construction of DTs that can restrain overfitting. If a classifier is to be constructed with a single DT only, d needs to be set with a not too small value. If DT is to be used as a weak learner of RF, it is recommended to set $d = 1$ to improve the individual weak learner's variance [13].

If these processes are executed with B sub-learning datasets, B-DTs are generated when feature values are input, and B classification results are output, which we call RF.

In constructing DTs as weak learners of RF, if you treat all feature values as solution candidates for the optimization problem in Eq. (11), the B weak learners generated will resemble each other. Majority decisions of resembling DTs would not help improve the generalization performance, which is derived from reduced variances. Therefore, in constructing DTs as weak learners of RF, we should search solutions from randomly selected γ feature values rather than all feature values. Thus, we can construct RF where weak learners have less similarity with each other. The recommended value of γ is the square root of the total number of feature values [13].

4.2. Principle for Importance Calculation

We next describe the principle for calculating the importance of feature values by applying the constructed classifier. We use OOB data that have not been used as sub-learning data. OOB data O_i is classified by weak learner $Tr(i)$. Classification errors acquired in the process, due to data unconcerned with learning, may correspond to the classifier's pseudo generalization performance. We call these classification errors as OOB Errors, where we can calculate the importance of feature values.

In one of the techniques for calculating the importance of feature values [13], increments in the classification er-

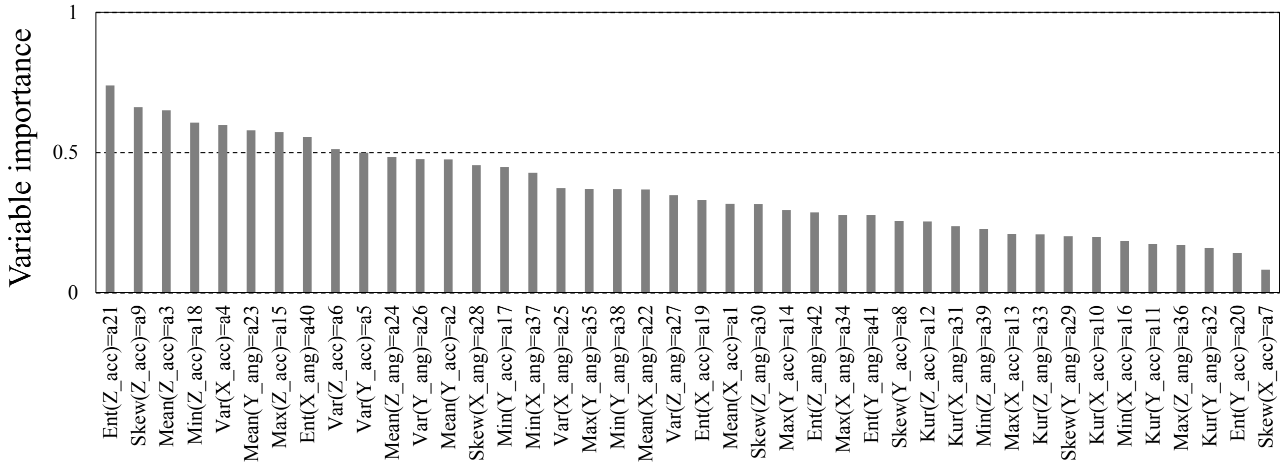

Fig. 3. Important feature values for the classification of four swimming styles (learning data).

rors when feature values a_j of OOB dataset O are randomly changed are taken as the importance I_j of a_j. The specific processes for deriving I_j are described below. First, randomly change the feature values a_j of the OOB data O_i that have not been used for the construction of weak learner $Tr(i)$. Then, obtain the weak learner $Tr(i)$ to classify the OOB data O_i and define any variations ΔE_j^i of the OOB Errors before and after the random changes of feature values as follows:

$$\Delta E_j^i = E_j^i - E^i, \quad \ldots \ldots \ldots \ldots \quad (13)$$

where E_j^i denotes the OOB Errors of the weak learner $Tr(i)$ when feature values a_j are randomly changed in O_i; E_i denotes OOB Errors of the weak learner $Tr(i)$. Next, calculate the average and variance of ΔE_j^i in all weak learner:

$$\overline{\Delta E_j} = \frac{1}{B}\sum_{i=1}^{B} \Delta E_j^i, \quad \ldots \ldots \ldots \ldots \quad (14)$$

$$\sigma_j^2 = \frac{1}{B}\sum_{i=1}^{B} (\Delta E_j^i - \overline{\Delta E_j})^2. \quad \ldots \ldots \ldots \quad (15)$$

Then, we define as the importance of feature values a_j:

$$I_j = \frac{\overline{\Delta E_j}}{\sigma_j}. \quad \ldots \ldots \ldots \ldots \ldots \quad (16)$$

This value becomes larger as the average $\overline{\Delta E_j}$ of increments in the classification errors of all weak learner is larger or as its standard deviation σ_j is smaller. Specifically, the larger and more uniform increments in the errors of each weak learner denote more important feature values.

4.3. Important Feature Values for the Classification of Four Swimming Syles: FV1

Using the technique described in Subsection 4.2, we have derived the feature values that are important for the classification of four swimming styles. Specifically, from learning data L, we have generated 100 sub-learning datasets $L^{\text{sub}} = \{L_i^{\text{sub}}\}_{i=1}^{100}$ by bootstrap sampling. In solving the optimization problem in Eq. (11), the number of feature values γ as solution candidates is the square root of the total number of feature values as recommended by [13] ($\gamma = \sqrt{42} = 6.48 \simeq 6$). We have also calculated the importance of feature values by applying the OOB dataset $O = \{O_i\}_{i=1}^{100}$ that have been generated incidental to the sub-learning dataset L^{sub}.

The results are shown in **Fig. 3**. The axis of the ordinate indicates the importance I_j of feature values, and the axis of the abscissa denotes each feature-value a_j. In this paper, we take feature values a_j with $I_j \geq 0.5$ as important feature values for the classification of four swimming styles. As a result, we have selected nine kinds of feature values: $a_3 = \text{Mean}(Z_{\text{acc}})$, $a_4 = \text{Var}(X_{\text{acc}})$, $a_6 = \text{Var}(Z_{\text{acc}})$, $a_9 = \text{Skew}(Z_{\text{acc}})$, $a_{15} = \text{Max}(Z_{\text{acc}})$, $a_{18} = \text{Min}(Z_{\text{acc}})$, $a_{21} = \text{Ent}(Z_{\text{acc}})$, $a_{23} = \text{Mean}(Y_{\text{ang}})$, $a_{40} = \text{Ent}(X_{\text{ang}})$. These feature values are hereinafter called FV1.

4.4. Important Feature Values for the Classification of Breaststroke and Butterfly: FV2

Previous researches on the classification of swimming styles have the problem of misclassifying breaststroke and butterfly. This problem could be solved by searching for feature values that are useful for classifying breaststroke and butterfly and by constructing a classifier based on these feature values. Therefore, in this research, besides FV1 described in Subsection 4.3, we select feature values specific to the classification of breaststroke and butterfly by first extracting just breaststroke and butterfly from the learning data L and then excluding feature values FV1. Then, we calculate the importance of the selected feature values in the same procedure as in Subsection 4.3.

The results are shown in **Fig. 4**, which highlights the importance of feature values just for classification of breaststroke and butterfly.

Too many feature values could make the feature-value space for classification of breaststroke and butterfly too

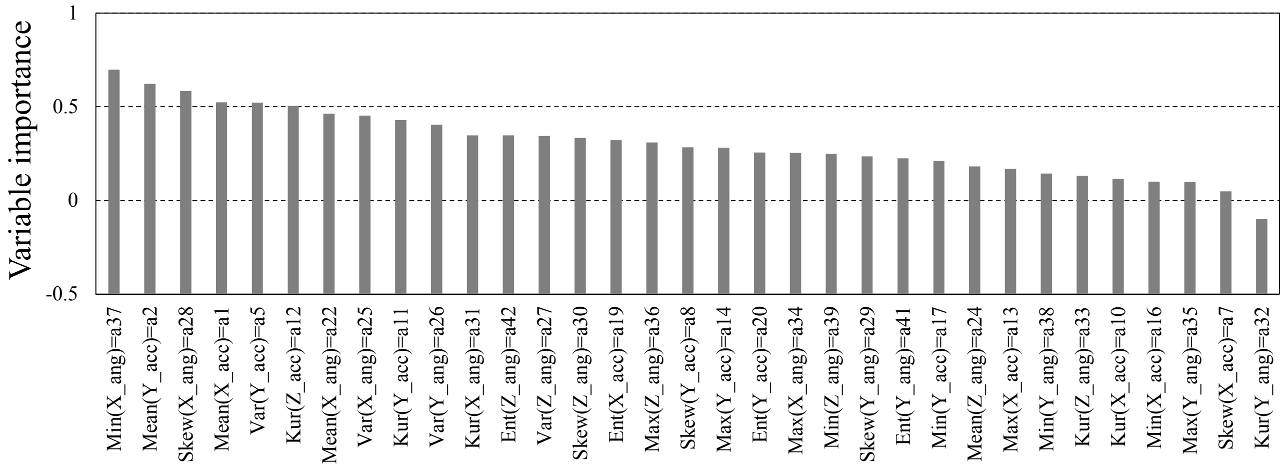

Fig. 4. Important feature values for the classification of breaststroke and butterfly (learning data).

broad, possibly increasing the classification errors for backstroke and front crawl. Therefore, in this paper, we have decided to adopt the top three kinds of feature values and have consequently selected the following three feature values: $a_2 = \mathrm{Mean}(Y_{\mathrm{acc}})$, $a_{28} = \mathrm{Skew}(X_{\mathrm{ang}})$, $a_{37} = \mathrm{Min}(X_{\mathrm{ang}})$. These feature values are hereinafter called FV2.

We have checked with the feature-value space as to whether or not the FV2 we have selected is effective for the classification of breaststroke and butterfly. We have used the learning data only. The results are shown in **Fig. 5**, where breaststroke is red circle and butterfly is purple box; butterfly on the upper and right side of the space and breaststroke on the lower and left side. Breaststroke and butterfly are comparatively clearly separated from each other, while breaststroke and butterfly are plotted close to each other. The use of FV2 alone may not be able to classify the two swimming styles with high accuracy, but a combined use of FV1, which is important for the classification of four swimming styles with FV2, should be able to improve the generalization performance for breaststroke and butterfly.

5. Construction of Swimming Style Classifier and its Evaluation

This research aims to solve the problem with the previous researches of mixing up breaststroke and butterfly in their classification by introducing (1) ensemble learning and (2) feature values specific to the classification of breaststroke and butterfly as well as constructing a swimming style classifier that can properly classify four swimming styles. To verify how effectively the introduction of (1) and (2) could solve this problem with the previous researches, we have constructed different swimming style classifiers where ensemble learning is applied or not applied and feature values specific to breaststroke and butterfly are applied or not applied and have compared their generalization performance. We have decided to take the

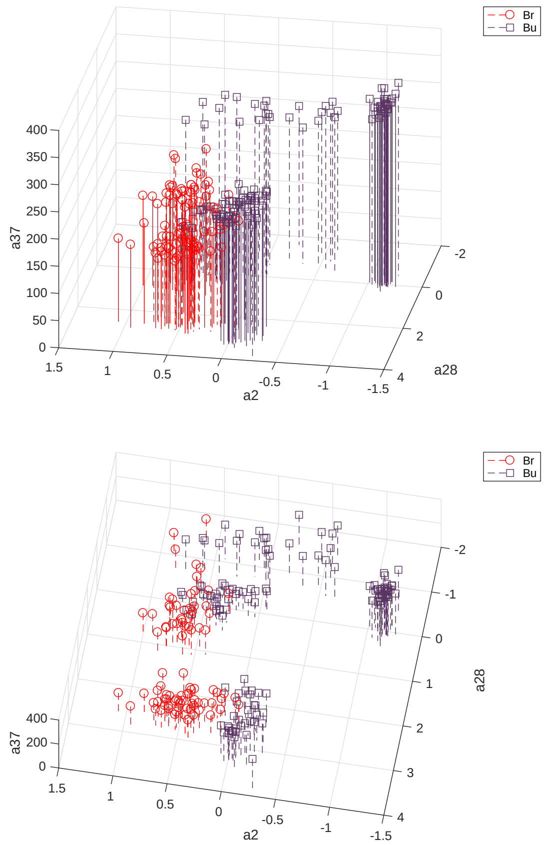

Fig. 5. Feature-value space of FV2 (Learning data).

classifier model with no ensemble learning as DT that is used by many of the previous researches on the classification of swimming styles [3–5]. The classifier model with ensemble learning as RF with DT is a weak learner.

5.1. Construction of Classifiers

The classifiers constructed in this research and their parameters are shown in **Table 3**. C1 and C2 are classifiers with DT. C1 has introduced the feature values FV1, which

Table 3. Constructed classifiers and their parameters.

ID	Model	Feature value	d	γ	B
C1	DT	FV1	30	-	-
C2	DT	FV1, FV2	30	-	-
C3	RF	FV1	1	3	41
C4	RF	FV1, FV2	1	3	136

C4 is proposed method.

Table 4. Structures of classifiers C1 and C2.

ID	IF	THEN
C1	$a_4 \geq 17.042 \wedge a_3 \geq 1.493$	Fr
	$a_4 \geq 17.042 \wedge a_3 < 1.493$	Ba
	$a_4 < 17.042 \wedge a_{21} \geq 2.149$	Br
	$a_4 < 17.042 \wedge a_{21} < 2.149$	Bu
C2	$a_4 \geq 17.042 \wedge a_2 \geq 1.043$	Ba
	$a_4 \geq 17.042 \wedge a_2 < 1.043$	Fr
	$a_4 < 17.042 \wedge a_{21} \geq 2.149$	Br
	$a_4 < 17.042 \wedge a_{21} < 2.149$	Bu

are important for classifying four swimming styles, while C2 has introduced feature values FV2 specific to breaststroke and butterfly in addition to FV1. The constraint d defined by Eq. (12) is 30, which is half of the 60 samples for backstroke, the smallest in the learning data, so that each swimming style can be allowed to have two subspaces on the feature-value space. Parameters γ and B are left blank because they are only applicable to RF. The classification criteria of DT are shown in **Table 4**.

C1 is branched into (butterfly/breaststroke) and (backstroke/front crawl) by $a_4 = \mathrm{Var}(X_{\mathrm{acc}})$. Swimmers make bilaterally symmetrical strokes in (butterfly/breaststroke) and bilaterally asymmetrical strokes in (backstroke/front crawl). As the X-axial direction indicates lateral motions, the size of horizontal directional variance seems to have worked effectively to classify bilaterally symmetrical and asymmetrical swimming styles. Backstroke and front crawl are divided by $a_3 = \mathrm{Mean}(Z_{\mathrm{acc}})$. The swimmer's whole body faces the floor in front crawl and the ceiling in backstroke. Therefore, the gravitational acceleration sign given in the Z-axial direction is reversed. These differences seem to be reflected in a_3. Breaststroke and butterfly are divided by $a_{21} = \mathrm{Ent}(Z_{\mathrm{acc}})$. The stroke waveforms of each swimming style, when represented by frequency domain, show differences in the size of fuzziness.

For C2, its classification criteria are almost the same as for C1. However, backstroke and front crawl are classified by $a_2 = \mathrm{Mean}(Y_{\mathrm{acc}})$. This suggests that accelerations in the direction of motion are different between backstroke and front crawl.

C3 and C4 are classifiers constructed by means of RF. C3 introduces feature values FV1 that are important for classifying four swimming styles, while C4 introduces feature values FV2 specific to the classification of breaststroke and butterfly in addition to FV1. For the constraint conditions d and the number of feature values γ to be

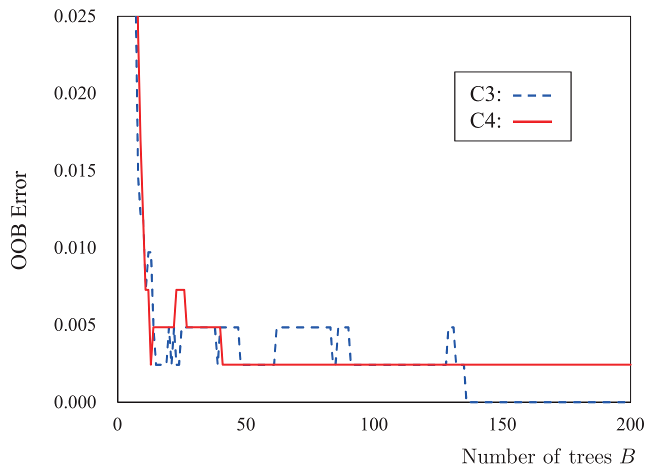

Fig. 6. Relation between the number of weak learners B and OOB Errors.

Table 5. Classification accuracies of learning data.

Classifiers C1 and C2									
		Measured				Evaluated			
	k	Ba	Br	Bu	Fr	A_k	P_k	R_k	F_k
Estimated results	Ba	60	0	0	0	1	1	1	1
	Br	0	89	9	0	.976	.908	.989	.947
	Bu	0	1	117	0	.976	.992	.929	.959
	Fr	0	0	0	136	1	1	1	1
Classifiers C3 and C4									
		Measured				Evaluated			
	k	Ba	Br	Bu	Fr	A_k	P_k	R_k	F_k
Estimated results	Ba	60	0	0	0	1	1	1	1
	Br	0	90	0	0	1	1	1	1
	Bu	0	0	126	0	1	1	1	1
	Fr	0	0	0	136	1	1	1	1

Remarks: Classification results of C1 and C2 are same with each other and those of C3 and C4 are also same with each other.

randomly selected in solving the optimization problem, we have adopted the recommended values [13]: 1 for d and the square root of the total number of feature values for γ. As FV1 consists of nine feature values, we have adopted $\gamma = \sqrt{9} = 3$ for C3. As the total number of feature values of FV1 and FV2 is 12, we have adopted $\gamma = \sqrt{12} = 3.46 \simeq 3$ for C4. For the total number of weak learners B, we have varied it between 1 and 200 to adopt a value where OOB Errors get converged. The results are shown in **Fig. 6**. As C3, OOB Errors get converged when $B \geq 41$. we have adopted 41 as B for C3. As C4, OOB Errors get converged when $B \geq 136$, we have adopted 136 as B for C4.

5.2. Comparison of Classification Accuracies

5.2.1. Classification Accuracies of Learning Data

Table 5 shows the classification results of learning data by the four classifiers constructed in Subsection 5.1. As C1 and C2 had the same classification results of learning data, the classification accuracies of C1 and C2 are given together, similar with those of C3 and C4. The diagonal lines in table represent the number of correctly classified swimming styles, and the other represent the number of erroneously classified swimming styles. TP_k denotes the number of correct classifications of the swimming style k as swimming style k. TN_k denotes the number of correct classifications of swimming styles other than swimming style k as other than swimming style k. FP_k denotes the number of erroneous classifications of swimming styles other than swimming style k as swimming style k. FN_k denotes the number of erroneous classifications of swimming style k as swimming styles other than swimming style k. We have calculated the accuracy A_k, precision P_k, recall R_k, and F-measure F_k for the classifications of the swimming style k as follows:

$$A_k = \frac{TP_k + TN_k}{TP_k + TN_k + FP_k + FN_k}, \quad \ldots\ldots (17)$$

$$P_k = \frac{TP_k}{TP_k + FP_k}, \quad \ldots\ldots (18)$$

$$R_k = \frac{TP_k}{TP_k + FN_k}, \quad \ldots\ldots (19)$$

$$F_k = \frac{2P_kR_k}{P_k + R_k}. \quad \ldots\ldots (20)$$

The evaluation results show that for DT, almost all swimming styles have been correctly classified except for a few wrong answers about breaststroke and butterfly. For RT, all swimming styles have been correctly classified. They just represent classification accuracies of learning data and do not guarantee classification accuracies of unknown data. Therefore, we verify in the following subsection the classifier's generalization performance for data that are not involved in learning.

5.2.2. Classification Accuracies of Test Data

The classifier we construct in this research is aimed at being able to correctly classify unknown third party's swimming styles rather than learning data. In this subsection, we measure the classifier's generalization performance from this perspective.

In general machine learning, the available techniques to measure generalization performance include n-hold Cross Validation (n-CV) and Leave One Out Cross Validation (LOO-CV) [13]. In the n-CV, generalization performance is measured by dividing all feature-value vectors into n datasets: $n-1$ datasets as learning data and one dataset as test data. Then, alter the dataset to be assigned as test data in n times and calculate the mean value of generalization performance. In the COO-CV, extract one feature-value vector out of all feature-value vectors as test data, where generalization performance is measured. Alter the feature-value vector to be extracted as test data for all of the feature-value vectors, and calculate the mean value of generalization performance.

These techniques of measuring generalization performance are particularly recommended in the situations where the number of available data is not sufficiently large [13]. If we adopt these techniques in this research, it could result in too high generalization performance. In this research, we calculate the generalization performance by the sliding window technique, where feature-value vectors overlap each other. Feature-value vectors calculated at certain points in the sensor data waveforms should partially retain numerical information on the feature-value vectors calculated before and after the said points. If we apply the n-CV or LOO-CV to such structured data, feature-value vectors not completely independent of each other should be assigned to learning data and test data, which would make the n-CV or LOO-CV generate unintentionally higher generalization performance. In the n-CV or LOO-CV, which acquires the feature-value vectors for one subject's strokes, the feature-value vectors of the same subject would be divided into learning data and test data. As both learning data and test data refer to the same subject, the generalization performance measured on these data cannot be applied to unknown third parties. We have decided to not use n-CV or LOO-CV, but to measure the generalization performance of the four classifiers constructed in Subsection 5.1 on the test data of a subject other than the ones for learning data.

The classification results of the test data are shown in **Table 6**. We first refer to classifiers C1 and C2 with DT. Their classification results for backstroke and front crawl are good, while they have mixed up breaststroke and but-

Table 6. Classification accuracies for test data.

		Measured				Evaluated			
	k	Ba	Br	Bu	Fr	A_k	P_k	R_k	F_k
Classifier C1 (DT/FV1)									
Estimated results	Ba	31	0	0	0	.976	1	.816	.899
	Br	7	49	6	0	.878	.790	.681	.731
	Bu	0	23	78	0	.901	.772	.929	.843
	Fr	0	0	0	100	1	1	1	1
Classifier C2 (DT/FV1+FV2)									
Estimated results	Ba	31	0	0	5	.959	.861	.816	.838
	Br	7	49	6	0	.878	.790	.681	.731
	Bu	0	23	78	0	.901	.772	.929	.843
	Fr	0	0	0	95	.983	1	.950	.974
Classifier C3 (RF/FV1)									
Estimated results	Ba	38	0	0	0	1	1	1	1
	Br	0	66	19	0	.915	.776	.917	.841
	Bu	0	6	65	0	.915	.915	.774	.839
	Fr	0	0	0	100	1	1	1	1
Classifier C4 (RF/FV1+FV2), proposed method									
Estimated results	Ba	38	0	0	0	1	1	1	1
	Br	0	67	1	0	.980	.985	.931	.957
	Bu	0	5	83	0	.980	.943	.988	.965
	Fr	0	0	0	100	1	1	1	1

terfly similar to the previous researches. **Table 4** shows that feature values specific to breaststroke and butterfly (FV2) are not applied to classifier C2 particularly for the classification of breaststroke and butterfly, which seems to have resulted in the failure of classifier C2 in resolving the mixed-up classification of breaststroke and butterfly despite the introduction of feature values FV2, which are effective for the classification of breaststroke and butterfly. Furthermore, the comparison in F_k of backstroke and front crawl between C1 and C2 confirms that C2 with feature values specific to breaststroke and butterfly FV2 introduced has lower values than C1. We review this matter on the branch structures of C1 and C2 shown in **Table 6**. The difference between C1 and C2 only lies in that C1 uses feature values a_3 to classify backstroke and front crawl and C2 uses the feature value a_2, which represents feature values specific to breaststroke and butterfly (**Fig. 4**), to classify backstroke and front crawl. This improper use of feature values seems to have reduced C2's F-measure of test data. The DT construction algorithm elects feature values effective for classification by applying the optimization problem expressed by Eq. (11) to learning data. Specifically, it interprets feature values to classify learning data most accurately as important feature values. The classification accuracies of learning data by C1 and C2 (**Table 5**) shows that both of them have correctly classified backstroke and front crawl. Feature values a_2 should be extremely important feature values from the viewpoint of maximizing the classification accuracy of learning data, so that feature values a_2, which are important for the classification of breaststroke and butterfly, seem to have been used for the classification of backstroke and front crawl. As learning data and test data are concerned with different subjects, their results plotted on the feature-value space are different from each other. The use of a_2 or a_3 has correctly classified all learning data, but it is only C2, which uses a_2, that has mixed up the test data on backstroke and front crawl. The feature-value selection algorithm introduced in DT, which aims to classify correctly all learning data, seems to have caused the results of C2.

In the process of constructing C1 and C2, the low-dimensional algorithm introduced in DT has searched for effective feature values to the feature values selected by the feature-value importance in **Figs. 3** and **4**. Specifically, C1 and C2 each have two feature-value selection processes. In the former process, learning data have pseudo test data (OOB data) and interpreted as important feature values those feature values that are more likely to have some effects on errors. The process introduced in DT is interpreted as important feature values for feature values that maximize the classification accuracies of learning data themselves. With this difference in the way of thinking about the "quality of feature values," it is possible that feature values that had been treated as having high generalization performance in the former process would be rejected in the latter process. This scenario would explain why some feature values important for the classification of four swimming styles (for example, a_9 and a_{18}) in **Fig. 3** were not used in the classifiers C1 and C2. Among the numerous classifier construction techniques, this is a phenomenon that could only occur in DT, which classifies feature values of learning data on minimum necessary feature values. Although classifiers with RF are also based on DT as described below, they divide learning data into countless sub-learning datasets or treat the data as solution candidates for γ feature values that are randomly selected in solving the optimization problem expressed by Eq. (11). Therefore, they can use not just the minimum necessary, but as many other feature values as possible.

Next, we refer to classifier C3 with RF and FV1 as well. **Table 6** shows the improved classification results of backstroke and front crawl. It slightly mixes up breaststroke and butterfly, but much less than C1 and C2. This seems to suggest that the introduction of ensemble learning is effective to improve the classifier's generalization performance for all swimming styles.

Lastly, we refer to classifier C4 with RF and FV1 + FV2 as well. Ensemble learning as well as feature values specific to breaststroke and butterfly are introduced into C4. This is a new swimming-style classifier that we propose in this paper. **Table 6** shows that C4 has least mixed up with breaststroke and butterfly among all classifiers or has properly classified the swimming styles. Feature values FV2 are specific to the classification of breaststroke and butterfly and are not so important for the classification of backstroke and front crawl. We are at first concerned that its expanded feature-value space may increase its wrong answer rates for backstroke and front crawl, but its actual classification results are found similar with classifier C3 with no FV2 introduced yet. This seems to suggest that feature values specific to the classification of breaststroke and butterfly (FV2) do not affect the classification of other swimming styles in RF; therefore, it would rather be effective to eliminate mixing up breaststroke and butterfly. F-measure for all swimming styles are ≥ 0.950, proving that it has high generalization performance.

5.2.3. Effects of Ensemble Learning and Proposed Feature Values

We quantitatively evaluate how much (1) ensemble learning and (2) feature values specific to the classification of breaststroke and butterfly, which we have introduced in this research, have actually contributed to improvements in classification accuracies. We use F-measure F_k as evaluation criteria. F-measure of each classifier for test data are shown in **Table 7**, where F-measure for four swimming styles, their mean values (Mean(ALL)), and mean values for breaststroke and butterfly (Mean(Br, Bu)) are shown.

We first refer to the effects of the introduction of ensemble learning. We have acquired these effects by deducting the F-measure of classifier C1 from the F-measure of classifier C3. The effects acquired in that way show an improvement of $+0.052$ in the mean value for all swimming styles and an improvement of $+0.053$ in Mean(Br, Bu).

Table 7. Comparison of F-measure of classifiers (test data).

Classifier	Model	Feature value	Ba	Br	Bu	Fr	Mean(ALL)	Mean(Br, Bu)
C1	DT	FV1	.899	.731	.843	1	.868	.787
C2	DT	FV1+FV2	.838	.731	.843	.974	.847	.787
C3	RF	FV1	1	.841	.839	1	.920	.840
C4 (proposed method)	RF	FV1+FV2	1	.957	.965	1	.981	.961

Next, we refer to the effects of the introduction of feature values specific to breaststroke and butterfly (FV2). We have acquired these effects by deducting F-measure of classifier C3 from the F-measure of classifier C4. The effects acquired in that way show an improvement of $+0.121$ in Mean(Br, Bu).

Lastly, we refer to the comprehensive effects of the introductions of ensemble learning and FV2. We have acquired the said effects by deducting F-measure of classifier C1 from F-measure of classifier C4. The comprehensive effects acquired in that way show an improvement of $+0.174$ in Mean(Br, Bu).

These results prove that the introductions of (1) ensemble learning and (2) feature values specific to breaststroke and butterfly (FV2) in this paper are found to be effective in resolving the problem with the previous researches of mixing up breaststroke and butterfly.

Table 8. Classification accuracies of classifier C5.

Learning data									
		Measured				Evaluated			
	k	Ba	Br	Bu	Fr	A_k	P_k	R_k	F_k
Estimated results	Ba	60	0	0	0	1	1	1	1
	Br	0	90	0	0	1	1	1	1
	Bu	0	0	126	0	1	1	1	1
	Fr	0	0	0	136	1	1	1	1
Test data									
		Measured				Measured			
	k	Ba	Br	Bu	Fr	A_k	P_k	R_k	F_k
Estimated results	Ba	38	0	0	0	1	1	1	1
	Br	0	59	0	0	.956	1	.819	.901
	Bu	0	13	84	0	.956	.866	1	.928
	Fr	0	0	0	100	1	1	1	1

5.2.4. Dimensional Effects of FV2

We have confirmed in the discussions that the introduction of feature values specific to breaststroke and butterfly (FV2) has reduced erroneous classifications of breaststroke and butterfly. These improvements in classification accuracies could be attributed to higher dimension of the feature value space as classifier C4 with FV2 introduced is 12-dimension, while classifier C3 is 9-dimension. To check in detail whether or not FV2 is really important for the classification of breaststroke and butterfly, we need to construct a classifier in the same dimensions with C3 on the feature value space where FV2 is introduced and to compare their classification accuracies with each other.

First, we have eliminated three feature values (a_6, a_{15}, a_{40}) of lowest importance from among nine feature values important for the classification of four swimming styles. We have then added three feature values important for the classification of breaststroke and butterfly (FV2) to make a total of nine feature values, which we call FV3.

Next, we have constructed a four-swimming-style classifier with RF using FV3, which we call C5. We have evaluated classifier C5's classification accuracies with learning data and test data. The evaluation results are shown in **Table 8**. Classifier C5 has correctly classified all the swimming styles in learning data similar to C3 and C4. However, it has 13 cases of misclassifying breaststroke for butterfly with test data. As it has no other misclassifications, it has generally good classification accuracies.

To check the effects of FV2 and higher dimension, we have compared the C5's classification results with F-measure of C3 and C4 as similar models constructed with RF. The comparison results are shown in **Table 9**. Comparison of mean F-measure of breaststroke and butterfly (Mean(Br, Bu)) between C3 and C5 shows a difference of $+0.074$. FV3 represents these feature values, where three feature values are first eliminated from FV1 and then FV2 is added instead. Given that C3 and C5 have the same dimensional feature value spaces, the improvement may be attributed to FV2. On the other hand, comparison of Mean(Br, Bu) between C4 and C5 shows a difference of $+0.047$. Both C4 and C5 have FV2 introduced in and the only difference between them lies in the dimensions of their feature value spaces. Therefore, the improvement may be attributed to an increase in the number of dimensions.

In summary of the discussions, classifier C4's high generalization performance of Mean(Br, Bu) $= 0.961$ may be attributed to the combined effects ($+0.074$) of the introduction of feature values specific to breaststroke and butterfly FV2 and the effects ($+0.047$) of higher dimension of the feature value space. The results seem to suggest that any even higher dimension of the feature value space will not lower the classifier's generalization performance. Therefore, we need to search for optimum dimensions of the feature value space to achieve higher generalization performance.

Table 9. Generalization performance for breaststroke and butterfly in RF.

Classifier	Feature Value	Model	Dimensions	Mean(Br, Bu)
C3	FV1	RF	9	.840
C4	FV1+FV2	RF	12	.961
C5	FV3	RF	9	.914

6. Evaluation Experiments by Simulations

6.1. Overview and Purpose

The discussions in Section 5 suggest that classifier C4 constructed in this research can properly classify all swimming styles of unknown third parties. However, the minimum value of classifier C4's precision rate P_k is 0.943, which indicates that C4 could misclassify swimming style once in twenty times. From the pattern recognition point of view, the precision rate should be determined as good result. From the user's point of view, however, we should aim to achieve much higher accuracies.

One of the possible techniques to achieve much higher accuracies would be to take a majority. In the process of converting feature values described in Subsection 3.2.3, data necessary to convert the first feature value is stored 1.08 s after the start of strokes. Then, data convertible into feature values are acquired every 0.54 s. Specifically, the data necessary for the first classification of swimming styles is collected 1.08 s after the start of strokes and then the data necessary for the subsequent classifications of swimming styles are sequentially accumulated every 0.54 s. This enables us to take a majority of classification results and adopt it to achieve much higher classification accuracies.

As mentioned at the first part of Section 1, the system [2] to be constructed in this research aims at feeding back swimmer's performance in real time to swimmers in a race. As high-speed classifications of swimming styles are required in the step prior to performance derivation, the system under development needs to be able to specify swimming styles at the early stage of a swimming race. If we want to specify swimming styles by a majority of classifier C4's classification results, the more classification results subject to a majority decision, the slower speeds to classify swimming styles. Therefore, the swimming-style classification should better be made by a majority with few classification results as possible.

In this section, therefore, we estimate by simulations on how many seconds after the start of strokes the system can classify swimming styles with high accuracy.

6.2. Preparation of Datasets

First, we prepare datasets for the simulation evaluations. We calculate that mean values and standard deviations of the feature values as acquired from test data on each swimming style and create a normal distribution of feature values for each swimming style. We generate from its normal random numbers 200 swims per swimming style (800 swims in total). Feature values of ten classifications for each swimming style are stored in data. Classifier C4 classifies swimming styles on such data to acquire ten classification results per swim. The first classification of swimming styles is made at 1.08 s after the start of strokes and then subsequent classifications are sequentially made every 0.54 s.

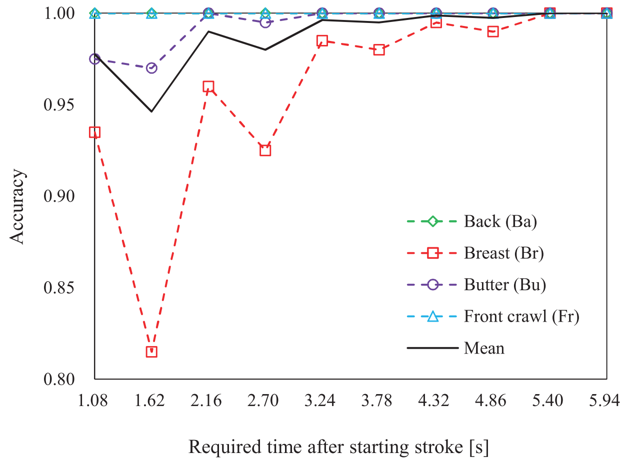

Fig. 7. Swimming style classification accuracies at elapsed time.

6.3. Results and Discussions

We have verified the classification accuracies by a majority of individual swimming-style classification results. The results are shown in **Fig. 7**. It shows the swimming-style classification accuracies as decided by a majority: the first one at 1.08 s after the start of strokes, the second one at 0.54 s later or at 1.62 s after the start of strokes, and subsequent classifications in the same way. Colored broken lines indicate the mean value of individual four swimming styles and a black solid line shows the mean values of four swimming styles.

For backstroke and front crawl, even if normal random numbers are added to test data that are not involved in learning, it shows they are all correctly classified (in **Fig. 7**, the axis of ordinate overlaps the accuracy line of 1.00). It tells that classifier C4 has very robust generalization performance for front crawl and backstroke.

For breaststroke and butterfly, the classification accuracies tend to converge at 100% as time elapses. This seems to be attributable to the decision by a majority of classifications where a few misclassifications if any would become negligible because of far more correct classifications. Mean classification accuracy for all swimming styles are 99.0% at 2.16 s after the start of strokes, 99.6% at 3.24 s, and 99.9% at 4.32 s and get better as time elapses, so that using the feature values to be acquired in the period of 5.4 s, it has obtained the correct answer rates of 100% for all of 800 swims. These results suggest that use of classifier C4 should enable us to collect data on which to classify properly swimming styles from 2.16 s to 5.4 s after the start of strokes. It also shows that the classification accuracies have dropped at 1.62 s after the start of strokes, probably because of the decision by a majority of the two classifications of swimming styles, where if the two classifications are not correct, their majority decision will turn out to be incorrect classifications as well.

The data on which we have derived the results are based on the test data that are not involved in the construction of classifiers. We have generated the data according to the normal random numbers of the feature values as acquired from the test data so that the data contain a few feature values that are apart by two standard deviations and three standard deviations. Classifier C4, which we have constructed in this study, has properly classified swimming styles even on such data, which seems to prove that the technique to classify swimming styles by a majority of classifier C4's classifications should have extremely robust and high generalization performance.

7. Conclusion

We have been engaged in the researches with a final goal of constructing a swimming competitive ability improvement support system using sensors [2]. To achieve the goal, swimming styles need to be automatically classified at a high speed and with accuracy after the start of a swimming competition. There are several previous researches available on automatic classifications of swimming styles. They are mainly based on non-ensemble learning and have a common problem of mixing up breaststroke and butterfly [3–6]. To resolve this problem, we have constructed a four-swimming-style classifier by introducing (1) ensemble learning and (2) feature values specific to breaststroke and butterfly.

Then, we have examined on test data how much the introductions of (1) and (2) above have really contributed to the improvements of the classifier's generalization performance. The examination results show that the introduction of (1) ensemble learning has improved the F-measure of classifications for all swimming styles by $+0.052$ and the F-measure of classifications for breaststroke and butterfly by $+0.053$. The introduction of feature values specific to breaststroke and butterfly has improved the F-measure of classifications for breaststroke and butterfly by $+0.121$. With F-measure for all swimming styles being ≥ 0.950, classifier C4, which we have constructed in this research, has proven to have high generalization performance.

Finally, we have conducted evaluation experiments by simulations to check how many seconds after the start of strokes it can classify swimming styles with as few errors as possible. We have found from the experiments that the use of data acquired from 2.16 s to 5.40 s after the start of strokes enables it to classify swimming styles with an accuracy of 99.0% to 100%.

The swimming style classification technique described in this paper has following three issues to be addressed in the future.

First Issue: to determine the width w of the more proper stroke period. In this paper, we have determined w by the function of the total sum of normal distributions of stroke period in the four swimming styles, which has resulted in adopting w that is biased to the stroke period of butterfly probably because the standard deviations of stroke period of butterfly are smaller than those of other swimming styles. Nevertheless, this does not pose a big problem to the swimming style classifier we have proposed in this paper, which has high generalization performance. To achieve much higher generalization performance, however, we may need to adopt w that is not biased to any particular swimming style by, for example, adopting mean stroke period of all swimming styles as w.

Second Issue: to use a proper number of feature values. In this paper, we have calculated the importance of feature values I_j for all feature values. We have defined that as for FV1, this feature values that have an importance ≥ 0.5 are important for the classifications of four swimming styles, and that as for FV2, top three feature values are important for the classifications of breaststroke and butterfly. These criteria for adopting feature values are in no event considered objective. To achieve much higher generalization performance, we should vary the criteria, observe how the generalization performance would change, and then determine the optimum criteria for adopting feature values accordingly.

Third Issue: to use as much data as possible. In this paper, we have conducted analyses using the data from 13 subjects. Use of test data prepared in this paper has achieved high-accuracy classifications of swimming styles. However, we cannot strongly claim that the same accuracy could be secured in classifications of unknown third parties' swimming styles. Therefore, we will need to collect as many learning data and test data as possible in the future.

As soon as these issues are resolved, we intend to detect the timing to start swimming to automatically detect the timing to start classifications of swimming styles. Then, we intend to promote solutions to issues required of the envisaged system [2] such as the construction of a per-stroke performance quantification technique.

Acknowledgements

This work was supported in part by Nagaoka University of Technology Presidential Research Grant (A). This work was also supported in part by JSPS Grant-in-Aid for Scientific Research (C) (Grant Nos. 16K06156 and 17K05437; T. Akiduki and H. Takahashi), by JSPS Grant-in-Aid for Young Scientists (B) (Grant No. 17K13179; Y. Omae, Grant No. 26800129; H. Takahashi).

References:

[1] MEXT, "Decision of the Target Sports on Summer Olympics," http://www.mext.go.jp/b_menu/houdou/27/04/1356716.htm, [Accessed February 8, 2016]

[2] Y. Omae, Y. Kon, K. Sakai, H. Takahashi, T. Akiduki, C. Miyazi, Y. Sakurai, N. Ezaki, and K. Nakai, "Supporting Environment to Coach Swimmer by Data-driven Approach," JSME Symp.: Sports engineering and Human Dynamics 2015, A-14, 2015.

[3] Y. Kon, Y. Omae, K. Sakai, H. Takahashi, T. Akiduki, C. Miyazi, Y. Sakurai, N. Ezaki, and K. Nakai, "Swimming Style Classification

for Developing System of Swimming Performance and Technique Evaluation," JSME Symp.: Sports engineering and Human Dynamics 2015, A-15, 2015.

[4] Y. Ohgi, K. Kaneda, and A. Takakura, "A swimming style prediction using the chest acceleration," Symp. on Sports and Human Dynamics 2012, pp. 98-103, 2012.

[5] 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 Conf. on Embedded Network Sensor Systems, pp. 76-90, 2014.

[6] U. Jensen, F. Prade, and B. M. Eskofier, "Classification of Kinematic Swimming Data with Emphasis on Resource Consumption," 2013 IEEE Int. Conf. Body Sensor Networks (BSN), pp. 1-5, 2013.

[7] P. Siirtola, P. Laurinen, J. Roning, and H. Kinnunen, "Efficient Accelerometer-Based Swimming Exercise Tracking," 2011 IEEE Symp. Computational Intelligence and Data Mining (CIDM), pp. 156-161, 2011.

[8] W. H. Liao, Z. X. Liao, and M. J. Liu, "Swimming Style Classification from Video Sequences," The 16th IPPR Conf. on Computer Vision, Graphics and Image Processing, pp. 226-233, 2003.

[9] Sports Sensing Co., LTD., "9 Axis Waterproof Type Wireless Motion Sensor," http://www.sports-sensing.com/products/motion/inertia/motionwp01.html, [Accessed February 9, 2016]

[10] Sony, Digital Video Camera HDR-CX720V, http://www.sony.jp/handycam/products/HDR-CX720V/spec.html, [Accessed February 17, 2016]

[11] L. Bao and S. S. Intille, "Activity Recognition from User-annotated Acceleration Data," Pervasive Computing, pp. 1-17, 2004.

[12] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data," AAAI, Vol.5, pp. 1541-1546, 2005.

[13] T. Hastie, R. Tibshirani, J. Friedman, M. Sugiyama, T. Ide, T. Kamishima, T. Kurita, and E. Maeda, "The Elements of Statistical Learning: Data Mining, Inference, and Prediction," Kyoritsu Publisher, 2014.

Name:
Yuto Omae

Affiliation:
Department of Electrical Engineering, National Institute of Technology, Tokyo College

Address:
1220-2 Kunugida, Hachioji, Tokyo 193-0942, Japan

Brief Biographical History:
2016.3 Dr. Eng. degree from Nagaoka University of Technology
2016.4-2017.3 Researcher, Japan Institute of Sports Sciences
2017.4- Assistant Professor, National Institute of Technology, Tokyo College

Main Works:
• "Method to Detect Change of Motivation to Enroll in University by Survey of Career Perceptions," J. of Japan Society for Fuzzy Theory and Intelligent Informatics, Vol.27, No.5, pp. 743-756, 2015.

Membership in Academic Societies:
• Japan Society for Fuzzy Theory and Intelligent Informatics (SOFT)
• The Institute of Electronics, Information and Communication Engineers (IEICE)
• Japan Society for Educational Technology (JSET)

Name:
Yoshihisa Kon

Affiliation:
Department of Information and Management Systems Engineering, Nagaoka University of Technology

Address:
1603-1 Kamitomioka, Nagaoka, Niigata 940-2188, Japan

Brief Biographical History:
2017.3 M.Eng. degree from Nagaoka University of Technology

Name:
Masahiro Kobayashi

Affiliation:
Department of Information and Management Systems Engineering, Nagaoka University of Technology

Address:
1603-1 Kamitomioka, Nagaoka, Niigata 940-2188, Japan

Brief Biographical History:
2017.4- Master's Student, Nagaoka University of Technology

Name:
Kazuki Sakai

Affiliation:
Department of Information Science and Control Engineering, Nagaoka University of Technology

Address:
1603-1 Kamitomioka, Nagaoka, Niigata 940-2188, Japan

Brief Biographical History:
2015.4- Ph.D. Student, Nagaoka University of Technology

Name:
Akira Shionoya

Affiliation:
Department of Information and Management Systems Engineering, Nagaoka University of Technology

Address:
1603-1 Kamitomioka, Nagaoka, Niigata 940-2188, Japan

Brief Biographical History:
2013.4- Professor, Nagaoka University of Technology

Name:
Hirotaka Takahashi

Affiliation:
Department of Information and Management Systems Engineering, Nagaoka University of Technology
Address:
1603-1 Kamitomioka, Nagaoka, Niigata 940-2188, Japan
Brief Biographical History:
2013.4- Associate Professor, Nagaoka University of Technology

Name:
Takuma Akiduki

Affiliation:
Department of Mechanical Engineering, Toyohashi University of Technology
Address:
1-1 Hibarigaoka, Tempaku, Toyohashi, Aichi 441-8580, Japan
Brief Biographical History:
2015.4- Assistant Professor, Toyohashi University of Technology

Name:
Kazufumi Nakai

Affiliation:
Information and Control Engineering Department, National Institute of Technology, Toba College
Address:
1-1 Ikegamicho, Toba, Mie 517-8501, Japan
Brief Biographical History:
2017.4- Associate Professor, National Institute of Technology, Toba College

Name:
Nobuo Ezaki

Affiliation:
Information and Control Engineering Department, National Institute of Technology, Toba College
Address:
1-1 Ikegamicho, Toba, Mie 517-8501, Japan
Brief Biographical History:
2015.4- Professor, National Institute of Technology, Toba College

Name:
Yoshihisa Sakurai

Affiliation:
Sports Sensing Co., LTD.
Address:
2-2-1-403 Mukaino, Minami, Fukuoka 815-0035, Japan
Brief Biographical History:
2016- Engineer, Sports Sensing Co., LTD.

Name:
Chikara Miyaji

Affiliation:
Department of Creative Informatics, The University of Tokyo
Address:
7-3-1 Hongo, Bunkyo, Tokyo 113-8656, Japan
Brief Biographical History:
2016.4- Project Researcher, Department of Creative Informatics, The University of Tokyo