International Journal of Innovative Computing, Information and Control
Volume **15**, Number **2**, April **2019** ISSN 1349-4198
pp. 783–791

CONSIDERATION OF HUMAN MOTION'S INDIVIDUAL DIFFERENCES-BASED FEATURE SPACE EVALUATION FUNCTION FOR ANOMALY DETECTION

Masaya Mori[1], Yuto Omae[2], Takuma Akiduki[3] and Hirotaka Takahashi[1]

[1]Department of Information and Management Systems Engineering
Nagaoka University of Technology
1603-1 Kamitomioka, Nagaoka, Niigata 940-2188, Japan
s183368@stn.nagaokaut.ac.jp; hirotaka@kjs.nagaokaut.ac.jp

[2]Department of Electrical Engineering
National Institute of Technology, Tokyo College
1220-2 Kunugida, Hachioji, Tokyo 193-0942, Japan

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

Received June 2018; revised October 2018

Abstract. *There are many researches of various human activity recognitions from the data of inertial sensors by using machine learning. In these researches, it is important to consider the individual difference. Even if the subjects perform the same activity, the data obtained from each subject are of different behaviors. Thus, if we construct the feature space, there is a possibility that the human activity of each subject does not concentrate on one region. In this case, if we consider the anomaly detection of the human activities by using this space, it is difficult to draw the boundary between the normal and anomaly activities. Therefore, the evaluation index/function that can search better feature values/space for various people is necessary. In this paper, we propose an evaluation function named "Consideration of Human motion's Individual differences-based Feature Space (CHI-FS) evaluation function" for the anomaly detection. We also confirm the effectiveness of the CHI-FS evaluation function by using the simulation data and the data of inertial sensors during car driving.*
Keywords: Machine learning, Feature space, Human activity recognition, Anomaly detection, Mathematical optimization

1. **Introduction.** Recently, various human activity recognitions from the data of inertial sensors by using machine learning are investigated, e.g., Khan et al. [1] used a neural network to classify such motions as walking, running, sitting down and standing up; Lester et al. [2] classified the human activities such as walking, running, brushing teeth and riding an elevator with the Hidden Markov Model (HMM); He and Jin [3] classified the walking, running and jumping with the support vector machine; Ward et al. [4] used HMM to classify such motions as assembly work (using drills and vises); Omae et al. [5] classified the swimming styles (backstroke, breaststroke, butterfly and front crawl) with the random forest.

If we consider the human activity recognitions, the data of inertial sensors (such as acceleration and gyro data) are of different behaviors since there are individual differences in human activities by the height, body weight, habits, length of leg and so on. Therefore, if we investigate the various human activity recognitions by using machine

DOI: 10.24507/ijicic.15.02.783

learning, it is important to consider the individual differences and search the important feature values for various people [6]. To achieve it, the evaluation index/function that can search better feature values/space for various people is necessary. However, the existing index/evaluation function (e.g., between-class and within-class variance [7] and minimum reference set [6, 8]) cannot be considered for such individual differences. Khan et al. [1] and Ward et al. [4] selected the feature values by using the linear-discriminant analysis; Lester et al. [2] selected the feature values by using a modified version of AdaBoost proposed by [9]; He and Jin [3] performed the feature extraction by using the discrete cosine transform and the principal component analysis and composed the feature space. These methods also cannot be considered for individual differences.

The human activity recognition in machine learning is mainly classified into the classification and anomaly detection. Let us focus on the human activity recognition during car driving as an example. Figure 1 shows the schematic plot of each activity during car driving operations in each feature space, respectively. Each legend corresponds to each driving activity. Figure 1(a) shows the feature space that can classify between normal and other dangerous driving operations. Figure 1(b) shows the feature space that can classify normal, one-handed, distracted driving, use of smartphone and other anomaly driving, respectively. In Figure 1(c), it is difficult to draw the boundary of each operation because the plots of each driving operation are mixed. The anomaly detection discriminates between the normal and anomaly plots. Therefore, the desirable evaluation index/function is to search the feature space like Figure 1(a). On the other hand, the classification classifies each plot. Therefore, the desirable evaluation index/function is to search for the feature space like Figure 1(b).

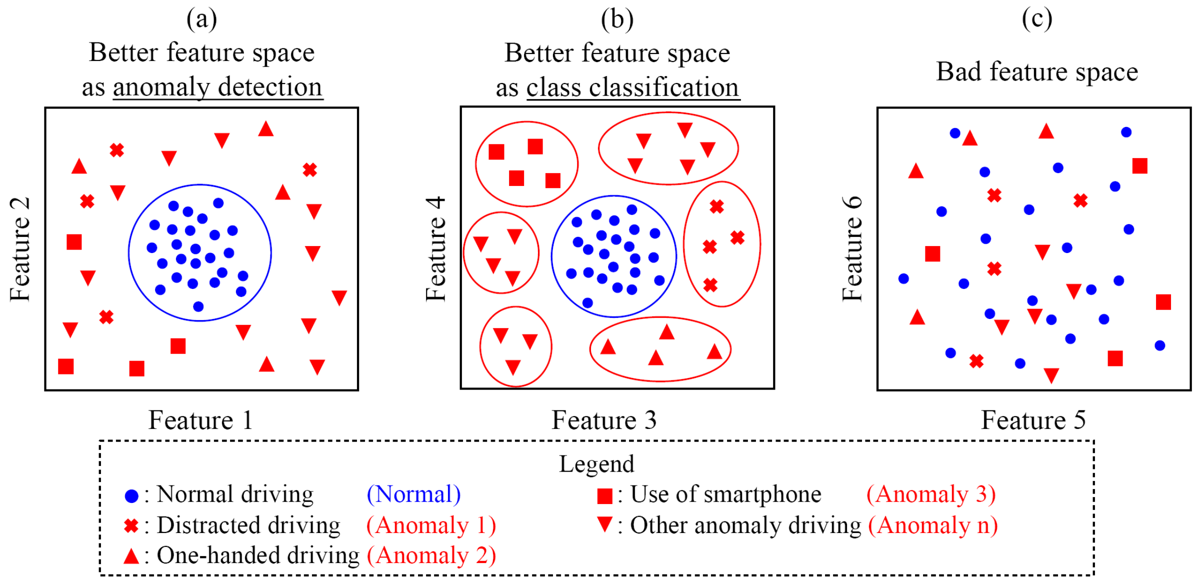

FIGURE 1. The difference of better feature space of anomaly detection and classification

The classification is highly convenient if the discriminated activity is determined beforehand. However, if there are various types of the anomaly activities such as Figure 1(b) and we cannot determine the discriminated activities beforehand, it is difficult to give the labels of all activities and obtain the learning data of all activities [10]. In addition, it is more relatively difficult to find the optimal space in the classification like Figure 1(b) than that in the anomaly detection like Figure 1(a). In most of the cases, we need to classify the normal and other dangerous activities.

Moreover, if we construct the feature space, there is a possibility that the human activity of each subject does not concentrate on one region. In this case, it is also difficult to draw the boundary between the normal and anomaly activities. Therefore, the evaluation index/function that can search better feature values/space for various people is necessary.

Therefore, in this paper, we propose an evaluation function named "Consideration of Human motion's Individual differences-based Feature Space (CHI-FS) evaluation function" for the anomaly detection.

The paper is organized as follows. In Section 2, we explain our proposed feature space evaluation function named "Consideration of Human motion's Individual differences-based Feature Space (CHI-FS) evaluation function". In Section 3, we confirm the effectiveness of the proposed CHI-FS evaluation function by using the simulation data and the data of inertial sensors during car driving. Section 4 is devoted to a summary.

2. Consideration of Human Motion's Individual Differences-Based Feature Space Evaluation Function.

2.1. Overview.

In this paper, since we consider the anomaly detection, we assume that the obtained data set from individual subject consist of many normal data and a few anomaly data.

The feature space composed of the feature values x is denoted by $\langle x \rangle$. We also define the following sets:

$$x_n, x_m \in X := \{x_1, x_2, \ldots, x_{F_{\max}}\}, \quad m, n = 1, \ldots, F_{\max}, \quad m \neq n, \tag{1}$$

$$i \in I := \{1, 2, \ldots, N_{\text{sub}}\}, \tag{2}$$

$$c \in C := \{\text{nor}, \text{ano}\}, \tag{3}$$

where X is the set whose elements are feature values and its size is $F_{\max}$, where $F_{\max}$ is the number of the considered feature values, I is the set whose elements are the subject's label, N_{sub} is the number of the considered subjects, and C is the set whose elements are the labels: normal or anomaly.

We aim to search for the effective/optimal feature space $\langle x^{\text{opt1}}, \ldots, x^{\text{opt}u} \rangle$ to judge normal and anomaly from the N_{sub} subjects data. To search for the optimal feature space, (i) we calculate the function to evaluate the effectiveness of the feature space for anomaly detection in each subject, (ii) we check the values of the function of all subjects and if the values of all subjects show uniformly effective then we adopt this feature space as the optimal feature space.

The overview of the construction of the evaluation function is shown in Figure 2. For simplicity, we consider $u = 2$ case in this paper and the feature space in Figure 2 can be denoted by $\langle x_n, x_m \rangle$. The method of the construction of the evaluation function is explained step by step.

2.2. Probability density function of the normal data.

It is natural that the normal data concentrate on the one region of the feature space like the top left panel of Figure 2 since the normal data consist of the data of the only specified activity. Therefore, we assume the multivariate normal distribution as the shape of the probability density function. Let the feature vector be $\boldsymbol{x} = (x_n\ x_m)^T$. The mean value of the feature value x_n, x_m in the subject i is defined as $\overline{x_n^i}$, $\overline{x_m^i}$ and the mean vector is denoted by $\boldsymbol{\mu}^i(x_n, x_m) = \left(\overline{x_n^i} \quad \overline{x_m^i}\right)^T$. The variance-covariance matrix of the normal data of the feature values x_n, x_m in subject i is defined as $V^i(x_n, x_m)$. Then the probability distribution of the normal data on the feature space $\langle x_n, x_m \rangle$ of subject i is given by:

$$P^i_{c=\mathrm{nor}}(x_n, x_m) = \frac{1}{2\pi\sqrt{|V^i(x_n, x_m)|}} \times \exp\left\{-\frac{1}{2}\left(\boldsymbol{x} - \boldsymbol{\mu}^i(x_n, x_m)\right)^T V^i(x_n, x_m)^{-1}\left(\boldsymbol{x} - \boldsymbol{\mu}^i(x_n, x_m)\right)\right\}. \quad (4)$$

From the probability distribution of Equation (4), we can estimate the distribution of the normal data of the subject i on the feature space, like the top middle panel of Figure 2.

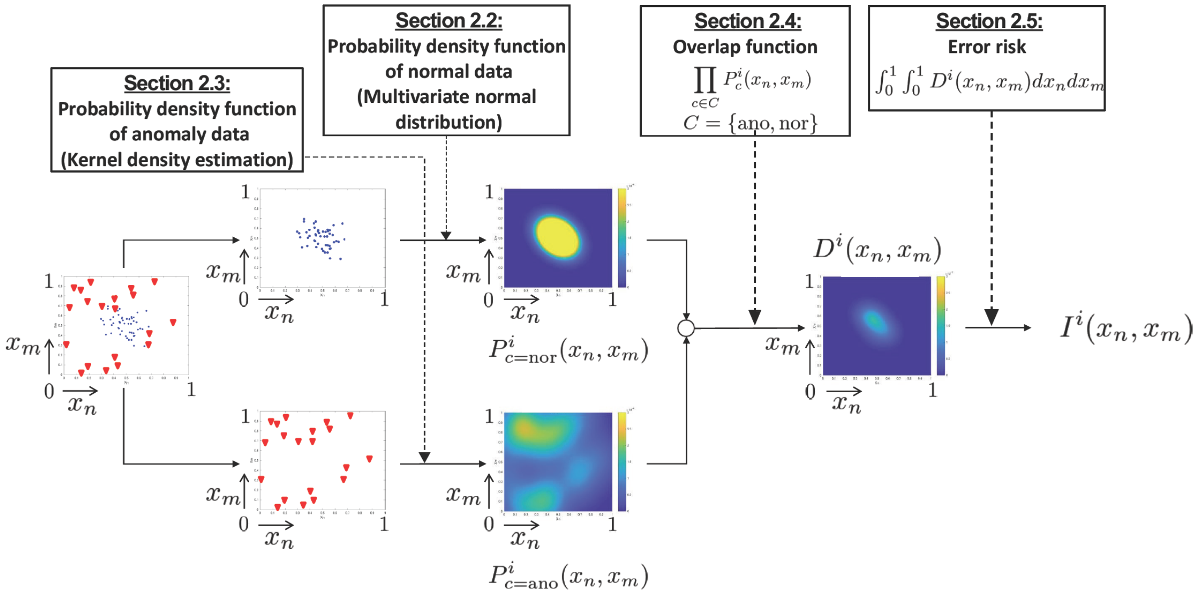

Figure 2. Overview of the construction of the evaluation function in feature space $\langle x_n, x_m\rangle$ of subject i (anomaly detection)

2.3. **Probability density function of the anomaly data.** The anomaly data do not concentrate on the one region of the feature space like the bottom left panel of Figure 2 since the anomaly data consist of the data deviated from the specified activity. Therefore, we assume the multivariate kernel distribution as the shape of probability density function.

The number of anomaly data in the feature values $\langle x_n, x_m\rangle$ in a certain subject i is N^i and the bandwidth in feature values $\langle x_n, x_m\rangle$ of subject i is defined as h^i_n and h^i_m, respectively [11]. Additionally, if the kernel function is used to the Gaussian kernel: $K(a) = \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}a^2\right\}$ and l-th anomaly data in feature values x_z of subject i is denoted as $x^i_{(z)l}$, then the probability distribution of the anomaly data on the feature space $\langle x_n, x_m\rangle$ of the subject i is given by:

$$P^i_{c=\mathrm{ano}}(x_n, x_m) = \frac{1}{N^i}\sum_{l=1}^{N^i}\prod_{z\in\{n,m\}}\frac{1}{h^i_z}K\left(\frac{x_z - x^i_{(z)l}}{h^i_z}\right). \quad (5)$$

From the probability distribution of Equation (5), we can estimate the distribution of anomaly data of the subject i on the feature space, like the bottom middle panel of Figure 2.

2.4. **Constitution of overlap function.** The evaluation function of the feature space is introduced by using the probability distribution of the normal and anomaly data.

In the feature space $\langle x_n, x_m\rangle$ of subject i, the probability distributions of the normal and anomaly data are compared. If the coordinate of the high probability of each feature space overlaps, it is natural that we consider this coordinate causes the misclassification. Thus, in each coordinate of the feature space, the following overlap function is calculated:

$$D^i(x_n, x_m) = \prod_{c \in C} P_c^i(x_n, x_m). \tag{6}$$

Note that it takes $0 \leq D^i(x_n, x_m) \leq 1$ by the definition of the probability density functions of Equations (4) and (5). Moreover, $D^i(x_n, x_m)$ takes a higher value if the coordinate of normal and/or anomaly is the higher probability of its occurrence. This schematic view is in the right middle panel of Figure 2.

2.5. **Error risk.** The error risk is defined as the function integrating the overlap function $D^i(x_n, x_m)$ of subject i:

$$I^i(x_n, x_m) = \int_0^1 \int_0^1 D^i(x_n, x_m)\ dx_n dx_m. \tag{7}$$

The error risk $I^i(x_n, x_m)$ corresponds to the volume of the overlap function $D^i(x_n, x_m)$. Being a one-dimensional real-valued function, it has a higher value if the coordinates of normal and anomaly with the higher probabilities of their occurrence overlap. Therefore, it is found that the feature space whose error risk is smaller is the important feature space for subject i.

2.6. **Optimization of the feature space.** Two optimal features $\langle x^{\text{opt1}},\ x^{\text{opt2}} \rangle$ for the anomaly detection can be obtained by solving the following optimization problem:

$$\langle x^{\text{opt1}}, x^{\text{opt2}} \rangle = \arg\min_{x_n, x_m} \left[\alpha \text{ mean}\{I(x_n, x_m)\}' + \beta \text{ std}\{I(x_n, x_m)\}'\right], \tag{8}$$

where mean$\{I(x_n, x_m)\}$ and std$\{I(x_n, x_m)\}$ represent the mean and standard deviation of the error risk of all subjects, respectively. The prime $'$ means the standardization. The standardization is defined as:

$$v\{I(x_n, x_m)\}' = \frac{v\{I(x_n, x_m)\} - \min(v\{I(x_n, x_m)\})}{\max[v\{I(x_n, x_m)\} - \min(v\{I(x_n, x_m)\})]}, \tag{9}$$

where v indicates the mean or std, and α and β represent the weight parameters which satisfy: $\alpha + \beta = 1$, $\alpha, \beta \in [0, 1]$. If the value of α is higher than 0.5, it treats the mean value of all subjects as important. In contrast, if β is higher than 0.5, it treats the standard deviation of all subjects as important. We call Equation (8) "Consideration of Human motion's Individual differences-based Feature Space (CHI-FS) evaluation function".

3. Evaluation of the Effectiveness of CHI-FS Evaluation Function.

3.1. Evaluation by using the simulation data.

3.1.1. *Outline.* To confirm the effectiveness of the CHI-FS evaluation function discussed in Section 2, we prepare the feature space of 14 types of simulation data assumed to be obtained from 5 subjects. We also prepare 4 patterns in the 2-dimensional feature space plot: 'Very Good', 'Good', 'Bad' and 'Very Bad', which mean the degree of mixing of the normal and anomaly data for each individual subject (the left side of Table 1).

The normal data are prepared of 50 plots by using the normal distribution:

$$\mathbf{u} = (0.5 \ \ 0.5)^T, \qquad S = \begin{pmatrix} 0.1^2 & 0 \\ 0 & 0.1^2 \end{pmatrix}, \tag{10}$$

where $\mathbf{u}$ is the mean vector of the normal data and S is the variance-covariance matrix of the normal data. The anomaly data is prepared 20 plots by using the uniformly random number.

TABLE 1. Evaluation results for feature space of each method

Feature space	Subjects 1	2	3	4	5	CHI-FS evaluation function $(\alpha, \beta) = (0.5, 0.5)$	CHI-FS evaluation function $(\alpha, \beta) = (0.7, 0.3)$	Between -class and within -class variance	MRS	Expected ranking
case1	Very Good	Very Good	Very Good	Very Good	Very Good	1	1	3	1	1
case2	Good	Good	Good	Good	Good	2	3	11	6	3
case3	Bad	Bad	Bad	Bad	Bad	6	7	8	14	
case4	Very Bad	Very Bad	Very Bad	Very Bad	Very Bad	9	12	12	7	
case5	Very Good	Very Good	Very Good	Good	Good	3	2	9	2	2
case6	Very Good	Very Good	Very Good	Bad	Bad	7	4	2	4	4
case7	Very Good	Very Good	Very Good	Very Bad	Very Bad	12	10	10	5	
case8	Very Bad	Very Bad	Very Bad	Very Good	Very Good	13	11	14	13	
case9	Very Bad	Very Bad	Very Bad	Good	Good	14	14	1	11	
case10	Very Bad	Very Bad	Very Bad	Bad	Bad	11	13	7	9	
case11	Good	Good	Good	Bad	Bad	4	5	13	3	5
case12	Bad	Bad	Bad	Good	Good	5	6	6	10	
case13	Very Good	Good	Bad	Very Bad	Good	10	9	5	8	
case14	Very Good	Good	Bad	Very Bad	Bad	8	8	4	12	

The 'Very Good' is the feature space satisfying the following condition: the anomaly data do not exist within 1σ region from the mean value of the normal data. This feature space can divide clearly between the normal and anomaly data. The 'Good' is feature space: 5 anomaly data exist within 1σ region from the mean value of the normal data. The 'Bad' is feature space: 10 anomaly data exist within 1σ region from the mean value of the normal data. The 'Very Bad' is feature space: 5 anomaly data exist within 1σ region from the mean value of the normal data. It is difficult to divide between the normal data and anomaly data in 'Very Bad' feature space.

The 14 cases obtained from 5 subjects we assumed are summarized in the left side of Table 1. The robust feature space means the feature space that can clearly distinguish between the normal and anomaly data of various subjects, i.e., case1, case2, case5. The case8 ('Very Bad' is 3 subjects, 'Very Good' is 2 subjects) is considered as the effective feature space for a few subjects. In contrast, the case11 ('Bad' is 2 subjects, 'Good' is 3 subjects) can be considered as the effective feature space for various subjects. Thus, in this paper, we should consider that case11 is also important. The expected top 5 ranking is shown in the right side of Table 1.

For comparison, the between-class and within-class variance [7] and Minimum Reference Set (MRS) [6, 8] are also used to evaluate the future space.

3.1.2. *Results and discussion.* The results obtained for the CHI-FS evaluation function, the between-class and within-class variance and Minimum Reference Set (MRS) are summarized in the right side of Table 1. The number from 1 to 14 represents the ranking in each evaluation method of the 14 feature spaces. The expected top 5 ranking is also shown in Table 1.

In the case of $(\alpha, \beta) = (0.5, 0.5)$ in the CHI-FS evaluation function of Equation (8), the evaluation results were: the first effective feature space was case1, the second one was case2 and the third one was case5. The case1 was the feature space prepared as the optimum, i.e., all are 'Very Good'. Thus, the desirable result was obtained. However, it was evaluated that case5 (3 'Very Good' and 2 'Good') was the third one and case2 (all 'Good') was the second one. This evaluation result was reversed in our expectation. In the case5, the mean of all subjects is better than that of case2. However, the distribution of the data is very deviated. As the result, the case2 ranking became higher than case5. From the point view of the proposed CHI-FS evaluation function of Equation (8), we confirmed that the second term tended to be emphasized. The same tendency was confirmed, e.g., between case4 and case7, between case4 and case8.

As one solution, we changed the weight parameters: increase the α and decrease the β, i.e., we considered the case of $\alpha > \beta$. In the case of $(\alpha, \beta) = (0.7, 0.3)$, the evaluation results are also shown in the right side of Table 1. We found that the desirable result was obtained in the case of $(\alpha, \beta) = (0.7, 0.3)$.

In the case of the between-class and within-class variance, the first effective feature space was case9, the second one was case6 and the third one was case1. The evaluation of the between-class and within-class variance is suited for class classification problem [7]. Thus, we found that this evaluation result was not desirable for our purpose.

In the case of MRS, the first effective feature space was case1, the second one was case5 and the third one was case11. Although the first and second one was obtained in the desirable result, the third one or later could not be obtained in the desirable result. For example, the case6 and case7 were not more robust than the case2, but the case6 and case7 have higher ranking than the case2.

These results suggest there is a possibility that we can search for the effective feature space considering the human motion's individual differences for anomaly detection by using the proposed CHI-FS evaluation function.

3.2. **Evaluation by using the data of the inertial sensors during car driving.**

3.2.1. *Outline.* In Section 3.1, we confirmed an effectiveness of the proposed CHI-FS evaluation function. In this section, we confirm more effectiveness of the proposed CHI-FS evaluation function by using the data of the inertial sensors during care driving [12]. Nagasawa et al. [12] tried to detect the aimless driving including the drowsy driving by using inertial sensors worn on the left and right wrists. By using these data and the proposed CHI-FS evaluation function, we search for the effective feature space to detect the aimless driving. From the inertial sensors worn on the left and right wrists, the data of X/Y/Z-axis acceleration, its composited acceleration and X/Y/Z-axis angular velocity are obtained with sampling frequency 100Hz. We set the 60sec window and 1sec slide width. To downsample the data, within 60sec window, we calculate the mean values of these data in 1sec sub-window. After that, the dynamic changes of these values are also calculated. The dynamic change represents the amount of the differential value from just before and current measurements. The mean, standard deviation, variance, skewness and kurtosis of the dynamic change within 60sec window are calculated. After that, we slide the window with 1sec slide width and repeat this procedure.

We use the mean, standard deviation, variance, kurtosis and skewness of the dynamic changes as the feature values. As the result, the 70 feature values are obtained. We also gave the label of normal and anomaly to the data. More details were explained in [12].

The number of the constructed feature spaces in this paper is ${}_{70}C_2 = 2415$. We evaluate and rank each feature space by using the CHI-FS evaluation function. Moreover, we evaluate the F-measure [12] as the anomaly detection quality of each subject by using the obtained each feature space and One Class Support Vector Machine (SVM).

3.2.2. *Results and discussion.* The results of the evaluation and the F-measure of each subject by using the obtained each feature space and One Class SVM are shown in Table 2. To confirm the effectiveness of CHI-FS evaluation function for various people, we calculate the mean and standard deviation of F-measure for all subjects. By definition of them, the large mean value of F-measure and the small standard deviation value of F-measure have to be selected at higher rank. We show the results of these values to the ninth line and tenth line of Table 2.

Table 2. Evaluation results by CHI-FS evaluation function and the quality of anomaly detection by using its space and One Class SVM

	CHI-FS evaluation value	Feature space*	F-measure (Individual subject)**					F-measure	
			Sub 1	Sub 2	Sub 3	Sub 4	Sub 5	Mean	Std
Evaluation value of the top 10 ranks	.0056 (1st rank)	(gy-r-sd, ax-l-ku)	**.846**	.076	.381	**.528**	**.485**	.463	.277
	.0060 (2nd rank)	(gx-r-sd, am-r-sd)	**.488**	**.572**	**.694**	.167	**.439**	.472	.196
	.0061 (3rd rank)	(am-r-sd, az-l-ku)	.241	**.515**	**.527**	.264	**.496**	.409	.143
	.0062 (4th rank)	(gx-r-sd, az-l-ku)	.258	.076	**.502**	**.426**	**.504**	.353	.184
	.0063 (5th rank)	(gx-r-sd, ax-l-ku)	**.789**	.133	**.574**	**.496**	**.538**	.506	.237
	.0063 (6th rank)	(az-l-sd, am-r-sd)	.318	**.615**	.324	.247	**.641**	.429	.185
	.0064 (7th rank)	(gx-r-sd, gz-l-sd)	.390	**.578**	.131	**.469**	**.503**	.414	.172
	.0069 (8th rank)	(am-r-sd, ax-l-ku)	.377	.199	**.510**	.340	**.508**	.387	.130
	.0070 (9th rank)	(gx-r-sd, ax-l-sk)	**.431**	.236	**.539**	.394	**.621**	.444	.147
	.0071 (10th rank)	(gx-r-sd, am-l-sd)	**.588**	.000	.231	**.711**	**.595**	.425	.298
Evaluation value ~ 0.10	.1001 (2061st rank)	(gx-r-sk, gy-l-sk)	.000	.047	.267	.288	**.549**	.230	.220
	.1002 (2062nd rank)	(ax-r-me, ax-r-sk)	.000	**.503**	.117	**.604**	**.446**	.334	.261
	.1006 (2063rd rank)	(gy-r-me, am-r-va)	**.417**	.016	**.414**	.076	.043	.193	.204
Evaluation value ~ 0.20	.2000 (2341st rank)	(gx-l-me, ay-l-va)	.000	.000	.006	.211	.310	.105	.146
	.2002 (2342nd rank)	(gx-l-me, am-r-ku)	.000	.000	.008	.168	.022	.040	.072
	.2031 (2343rd rank)	(gz-r-me, am-l-me)	**.491**	.016	.057	**.494**	.280	.267	.229
Evaluation value of the lowest ranks	.6251 (2413th rank)	(gx-l-me, gy-l-me)	.284	.000	.000	.146	.064	.099	.120
	.7340 (2414th rank)	(gz-r-me, gx-l-me)	.283	.031	.003	.088	.279	.137	.135
	.9316 (2415th rank)	(gx-l-me, gz-l-me)	.000	.000	.000	.123	.104	.045	.063

* 1st-digit: (ax/ay/az/am, x/y/z/composited acceleration) (gx/gy/gz, x/y/z angular velocity), 2nd-digit: (r/l, inertial sensor of the right/left wrist), 3rd-digit: (me/sd/va/sk/ku, mean/standard deviation/variance/skewness/kurtosis of the dynamic change)

** Boldface means F-measure $\geq .400$ and the underline means the maximum of F-measure in its space.

In case of the top 10 ranks, the mean value of F-measure for all subjects is $.353 \sim .506$, and the F-measure of each subject is large overall (Especially, the 4 out of 5 F-measures of the feature spaces of the 2nd and 5th rank exceed .400.). Moreover, about the standard deviation value of F-measure of all subjects, there are no extremely large dispersion in the feature space of top 10 ranks.

If the CHI-FS evaluation value becomes larger, then the mean of F-measure also tends to become worse. On the other hand, the standard deviation of F-measure is overall smaller dispersion, but this is because of extremely small F-measure of each subject if the evaluation value becomes larger. If the evaluation value of the CHI-FS evaluation function shows good value (small value), the detection accuracy for the normal and anomaly data is good. Note that even if the feature space is evaluated as the top rank, there are some cases that F-measure of subject 2 is low. There is a possibility that the subject 2 has the specific individual difference.

These results also suggest even if we use the real data of the inertial sensors, there is a possibility that we can search for the effective feature space considering the human motion's individual differences for anomaly detection by using the proposed CHI-FS evaluation function. However, we emphasize that we need more verification.

4. **Summary.** We proposed the feature space evaluation function named "Consideration of Human motion's Individual differences-based Feature Space (CHI-FS) evaluation function". We also confirmed the effectiveness of the proposed CHI-FS evaluation function by using the simulation data and the data of the inertial sensors during car driving. However, the validation is not sufficient because the number of subjects is low. We need more verification.

As the future work, we will perform more verification of the effectiveness and reliability to other case studies. In this paper, for simplicity, we discuss the case of the 2-dimensional feature space. The extension of Equation (8) for the n-dimensional feature space is not so difficult. However, it is easy to notice that the calculation cost becomes high. Thus, the reduction of calculation cost is also an important issue.

Acknowledgments. This work was supported in part by JSPS Grant-in-Aid for Young Scientists (B) (Grant No. JP17K13179; Y. Omae) and by JSPS Grant-in-Aid for Scientific Research (C) (Grant Nos. JP16K06156 and JP17K05437; T. Akiduki and H. Takahashi).

REFERENCES

[1] A. M. Khan, Y. K. Lee, S. Y. Lee and T. S. Kim, A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer, *IEEE Trans. Information Technology in Biomedicine*, vol.14, pp.1166-1172, 2010.

[2] J. Lester, T. Choudhury and G. Borriello, A practical approach to recognizing physical activities, *Int. Conf. on Pervasive Computing*, pp.1-16, 2006.

[3] Z. He and L. Jin, Activity recognition from acceleration data based on discrete consine transform and SVM, *IEEE Int. Conf. on Systems, Man and Cybernetics*, pp.5041-5044, 2009.

[4] J. A. Ward, P. Lukowicz, G. Troster and T. E. Starner, Activity recognition of assembly tasks using body-worn microphones and accelerometers, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.28, no.10, pp.1553-1567, 2006.

[5] 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.

[6] Y. Omae and H. Takahashi, Feature selection algorithm considered trial and individual differences for machine learning of human activity recognition, *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol.21, no.5, pp.813-824, 2017.

[7] M. Araki, *Speech Recognition Systems Made of Free Software: From Beginning of Pattern Recognition and Machine Learning to Dialog System*, Morikita Publication, 2007 (in Japanese).

[8] X. W. Chen and J. C. Jeong, Minimum reference set based feature selection for small sample classifications, *The 24th International Conference on Machine Learning*, pp.153-160, 2007.

[9] P. Viola and M. Jones, Rapid object detection using a boosted cascade of simple features, *The 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.511-518, 2001.

[10] M. Tada, F. Naya, M. Okada, H. Noma, T. Toriyama and K. Kogure, Detection of driver's anomaly behavior using wireless 3D-accelerometers, *Transactions of the Japanese Society for Artificial Intelligence: AI*, vol.23, pp.105-116, 2008.

[11] S. J. Sheather, Density estimation, *Statistical Science*, vol.19, no.4, pp.588-597, 2004.

[12] J. Nagasawa, T. Akiduki, I. Kamio, H. Takahashi, Y. Omae and Z. Zhang, A study of driver state estimation from physiological and behavioral signals, *The 60th Joint Automatic Control Conference (JACC) (Rengo 60)*, SuI1-1, 2017 (in Japanese).