Machine Learning of User Attentions in Sensor Data Visualization

Keita Fujino(✉), Sozo Inoue, and Tomohiro Shibata

Kyushu Institute of Technology, Fukuoka, Japan
fujino-keita@edu.brain.kyutech.ac.jp, sozo@mns.kyutech.ac.jp, tom@brain.kyutech.ac.jp

Abstract. In this paper, we propose a method for automatically estimating important points of large sensor data by collecting attention points of the user when visualized, and applying a supervised machine-learning algorithm. For large-scale sensor data, it is difficult to find important points simply through visualization, because such points are buried in a large scope of visualization. We also provide the results of an estimation, the accuracy of which was over 80% for multiple visualizations. In addition, the method has the advantage that the trained model can be reused to any other visualization from the same type of the sensors. We show the results of such reusability for the new type of visualization, which achieved an accuracy rate of 70–80%.

Keywords: Sensor data · Visualization · User attention
Attention points · Machine learning

1 Introduction

In recent years, the spread of smartphones and the miniaturization and price reduction of sensor devices have significantly progressed. Accordingly, various researches using sensors ongoing such devices are ongoing. Among them are researches on sensing behaviors at home and in hospital by using infra-red, temperature, and humidity sensors [1], and researches on behavioural sensing using position sensors and power consumption sensors [2]. As a method of utilizing the enormous amount of data obtained through these sensors, visualization is often conducted as the first step. Through visualization, important areas such as the location of an accident, or areas that need improvement for better business processes, can be found. In a hospital, for example, if any unusual movements that a nurse may make during rounds can be found using a mobile sensor or environmental sensor patterns, it becomes presumably possible to detect any abnormalities and how critical they are.

However, although such observations may be possible by visualizing small-scale data, it becomes difficult through human observation at a very large scale. Methods on automatically finding important points in a large sensor data are crucially required.

K. Murao et al. (Eds.): MobiCASE 2018, LNICST 240, pp. 125–143, 2018.
https://doi.org/10.1007/978-3-319-90740-6_8

In this paper, we assume that the important points to be visualized can be learned from users' visual attention, and propose a method for realizing appropriate visualization by applying a supervised machine-learning. In the proposed method, we firstly extract feature values from sequential sensor data, and train the detection model of important points with training labels of user's attention. Using the trained model, the important points to be visualized (which we call *attention points*) of any sensor data of the same type can be extracted.

Furthermore, because our method uses the feature vectors from sensors—instead of visualization as input, the method has the advantage that the trained model can be applied to any other visualization from the same type of the sensors.

To evaluate the proposed method, we used several types of sensor data collected from a nursing home, as well as the attention point labels collected from an experiment with recruited observers. As a result, for all of the sensor data, the accuracy reached over 80%. We also evaluated whether the algorithms learned using a single visualization method can be applied to other visualization methods. As a result, even when the learning algorithms were replaced with both an illumination sensor and an acceleration sensor, an estimation accuracy of 70–80% could be obtained.

The contributions of this paper are as follows.

1. We propose a method for extracting only the sensor data that the user is able to pay attention to from an enormous amount of data, and evaluated them, which showed high accuracies with real sensor data.
2. Unlike learning from images, we propose a method that is applicable to different visualization methods from the original visualization.
3. We made a prototype system for collecting the users' attention points. A dataset was created through the system by employing various observers.
4. We discussed the validity and usefulness of our method through an analysis of the variable importance using ensemble learning, and by showing the results of visualization.

2 Background

Utilizing the enormous amount of sensor data obtained by smartphones and sensor devices, users can find important information, such as the location of an accident or an item of interest through visualization. For example, in a hospital, if it can be determined from a mobile or environmental sensor that a nurse acted in an unusual manner, it is possible to detect the abnormalities of round targets and to assume how critical such events are. In addition, if the sensor records the fact that the staff are conducting too many movements when carrying out their duties and office work, it can be inferred that the room layout and work procedures are inefficient.

Although such observations may be possible by visualizing small-scale data, they become difficult even for human observation at a large-scale. For example, Fig. 1 visualizes the acceleration data a 15 h work period of a staff member at

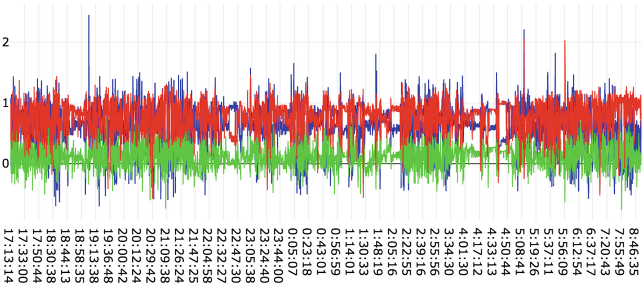


Fig. 1. An example of visualizing large-scale sensor data

a nursing facility. This figure does not reveal where individual work such as regular round was conducted, as well as what area we should focus on during an activity. Furthermore, when data becomes multivariate, manual search for optimal visualization may cause poor work efficiency. For this reason, technology to efficiently visualize sensor data and extract important areas from such data becomes very important.

With regard to such research, Wongsuphasawat et al. [7,8] proposed a system that automatically visualizes multiple objects and recommends a more useful visualization. This method, using several visualization methods, automates visualization by selecting the data variables, and recommends a useful visualization for the user. However, the system only presents various visualization methods and does not tell us where to put a focus. Therefore, it cannot be used for extracting useful sections from large data as shown in Fig. 1.

In addition, Walker et al. [9] proposed a system specialized for time series data. By specifying the attention points when observing the visualized time series data, the range can be enlarged and rendered. At this time, specified part is expressed in a tree. This allows an observation to be achieved while retaining the original information. Although such a visualization method can be easily displayed as long as it specifies attention points, there is no function for automatically detecting such points. Therefore, the problems of listing long-term data, such as in Fig. 1, and not knowing where to focus on remain.

As research on automating such attention points, a study by Bylinskii et al. [5] was considered. In addition, Bylinskii et al. proposed a method for estimating where to focus when viewing a graphic design or data visualization. Using a bitmap image in a graphic design or visualization data as input, the output is the attention points given as teaching data using click data, which can obtain the same result as gaze tracking, namely, a BubbleView [6], and learning and guessing using fully convolutional networks. As a result, the degree of importance is expressed as a heat map. As bitmap image of graphic design and visualization data as input, it makes output as an attention points given as training labels using click data which can obtain the same result as gaze tracking called BubbleView [6] by learning and guessing using Fully Convolutional Networks. However, when trying to use the method proposed by Bylinskii et al. for visualization of sensor data, when multiple visualizations are performed on the same sensor data, or when switching to another visualization method in the middle of the analysis, it is necessary to prepare training data again for the new visualizaion method. If multiple visualizations are made for one type of sensor data, it is possible to estimate the attention points without preparing new training data, and the cost of constructing the estimation model is kept low.

Based on above, the following system that estimates attention points for sensor data is crucially required.

1. When the observer's visual attention is given as the training data of the sensor data, points to be visualized can be automatically estimated as *an attention point* and presented to the new sensor data or a new observer.
2. By learning a certain visualization method, it is possible to automatically estimate the attention points even if another visualization method is applied to the same type of sensor data.

3 Proposed Method

In this section, we propose a supervised machine-learning approach based on sensor and training data of the attention points, and propose a method for estimating and presenting such points for new sensor data, and even for a new visualization method.

In the following, we first show the basic usage in estimating the attention points from the sensor data X when the estimation algorithm f is given, and applying the training of the estimation algorithm f.

3.1 Estimation

Let $X_{1:n}$ be an input sensor data sequence. $1 : n = (1, 2, \cdots, n)$ represents the sample number proportional to the timestamps. $X_i \in X_{1:n}$ can be either a scalar or a vector. The method estimates the attention points with these sensor data $X_{1:n}$ using the proposed method, and the set of attention points is (Fig. 2)

$$f(X_{1:n}) = \{(b_1, e_1), (b_2, e_2), \cdots, (b_m, e_m)\}, \tag{1}$$

where $1 < b_i, e_i < n$. Using this set of attention points $\{(b_j, e_j)\}_{j=1}^{m}$, parts of the sensor data

$$\{X_{b_j:e_j}\}_{j=1}^{m} \tag{2}$$

are extracted and displayed.

An example of $m = 2$ is shown in Fig. 3. The upper-half of the figure shows the sensor data X visualized. The lower-half of the figure shows the attention points estimated using the proposed method.

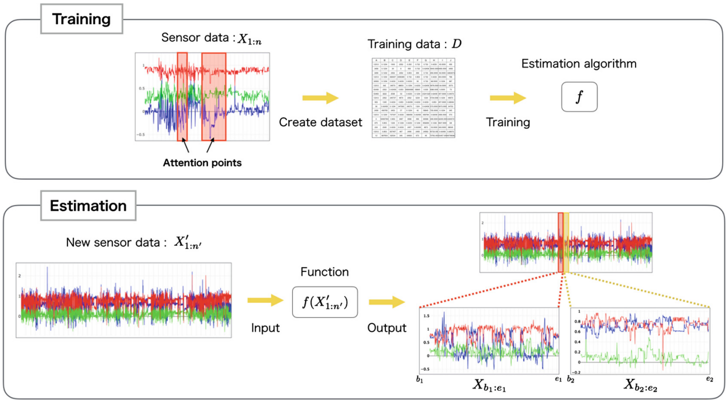


Fig. 2. Outline of the proposed method (At training, dataset D is created from the sensor data $X_{1:n}$ and the attention point set $\{Y_j\}_{j=1}^{m}$. From D, the estimation algorithm f is created by supervised machine-learning algorithm; At estimating, attention point set $\{Y'_j\}_{j=1}^{m'}$ is estimated by giving new sensor data $X'_{1:n'}$ as an input of the function f).

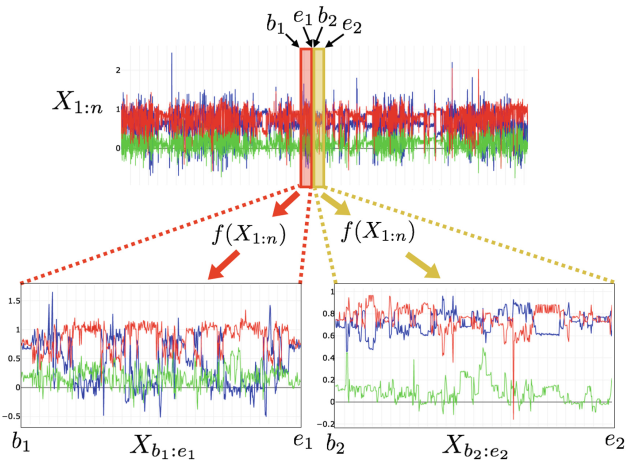


Fig. 3. An example of estimating attention points using the proposed method (upper half, visualizing sensor data X; lower half, attention points estimated using the proposed method).

3.2 Training

With the proposed method, we assume that the sensor data and the attention points of the viewers are collected. To obtain the training label of the attention points, we ask the user to observe the sensor series $X_{1:n}$, visualized as shown in Fig. 4, and to add a red frame to arbitrary areas where they focused their attention. At this time, the time at which the j-th red frame is attached is represented by the pair $y_j = (b_j, e_j)$ of the start point b_j and end point e_j. Actually, l visualizations are prepared for a plurality of sensor series $\{X^k_{1:n_k}\}^l_{k=1}$, and labeled for each visualization; for simplicity, however, only one series $X_{1:n}$ is described below.

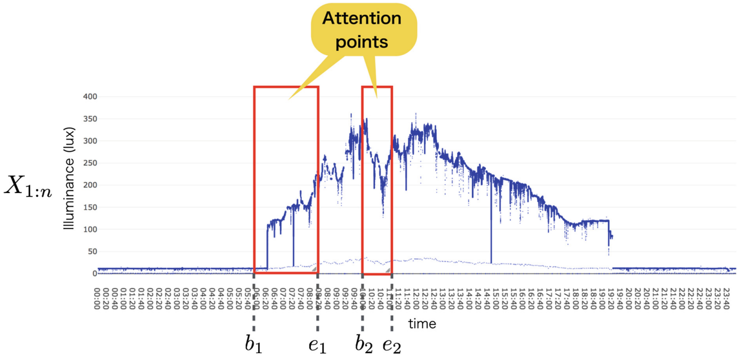


Fig. 4. An example of labeling attention points from visualized sensor data. The red frame indicates such points as labeled by the user. (Color figure online)

The procedure for training the attention points estimation algorithm from the dataset above is described below.

Algorithm

- Input: sensor series $X_{1:n}$, attention point set $\{Y_j\}^m_{j=1}$.
- Output: Function f to input a sensor series $X'_{1:n'}$ data, and output attention point set $\{Y'_j\}^{m'}_{j=1}$.

1. As shown in Fig. 5 B, sensor data are divided based on a fixed time width T, that is, a time window set $\{w(X_{1:n}, i)\}^N_{i=1}$ is obtained using a time windowing function $w(X_{1:n}, i) = ((i-1)T+1) : (iT)$. The margins are skipped for simplicity. For the sake of simplicity, we hereafter represent $w(X_{1:n}, i)$ as w_i.
2. For each time window w_i, a vector $V_i = h(X_{w_i})$ is calculated using a feature vector calculation function h using statistics and frequency components. Specific feature vectors are described in Sect. 4.3.
3. Compare the time window w_i and the attention points (b_j, e_j) of $b_j \in w_i$ or $e_j \in w_i$ in the time window for each j, which are collectively labeled as $Y_j = T$ if more than one-half of the samples in the time window are focused on, and as $Y_j = F$ otherwise. Apply this to all time windows $w_i, \cdots, w_N$, and obtain the training data $D = \{(V_j, Y_j)\}^N_{j=1}$ with N samples.
   Fig. 5 D shows an example of training data D. A blue row in the figure shows V_j, and a red row indicates the label attached as the attention points Y_j.
4. Apply supervised machine learning with the attention points Y_j as an objective variable, and the feature variable V_j as an explanatory variable, to obtain an estimated model g.

5. Output the function f, which calculates

$$\{\tilde{Y}_i\}_{i=1}^{N} = \{g \cdot h(X_{w_i})\}_{i=1}^{N} \tag{3}$$

first, and then convert $\{\tilde{Y}_i\}_{i=1}^{N}$ into $\{(b_j, e_j)\}_{j=1}^{m}$ by converting the jth changing time from F to T and then b_j, and jth time from T to F and then e_j in the order of i.

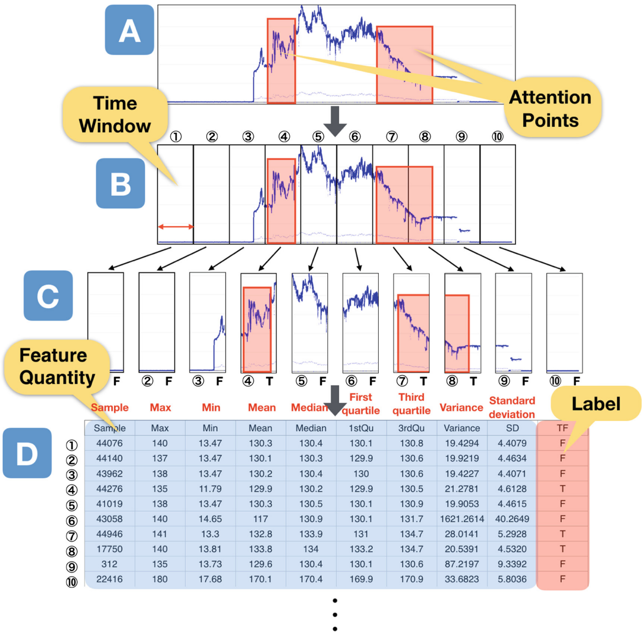


Fig. 5. Example of how to create a dataset. (A, the attention points labeled by the user; B, sensor data divided by a fixed time width T; C, calculation of the feature vector for each divided data, which are collectively labeled as $Y_j = T$ if more than one-half of the samples in the time window are focused upon, and as $Y_j = F$ otherwise; and D, example training data D. A blue row in the figure shows V_j, and a red row indicates a label attached as an attention point Y_j.) (Color figure online)

Using the function f obtained using the algorithm, the estimated attention points

$$\{(\tilde{b}'_i, \tilde{e}'_i)\}_{i=1}^{m'} = f(X'_{1:n'}) \tag{4}$$

are calculated for the new sensor series $X'_{1:n'}$.

This method is expected to be applicable to various visualization methods, because the feature vectors are calculated from the sensor data. Because Bylinskii et al. [5] create visualization data to be trained as an image, they do not expect to estimate the attention points from new visualization methods that have not yet been learned. With our method, because we use feature values from the sensor data, there is an advantage in that we can learn independently from visualization. Furthermore, if the sensor type is the same, any large-scale data can be handled.

4 Evaluation Experiment

In this section, to evaluate whether the attention points can be estimated using any visualization methods, we evaluate the following:

1. Can the proposed method estimate the attention points correctly?
2. Can we estimate for other visualization methods through the model trained by a certain visualization method?

Here, (1) is evaluated by the accuracy of the proposed algorithms. Regarding (2), we evaluate whether an estimation model trained by a certain visualization method can estimate correct attention points for other visualization methods for the same sensor type.

4.1 Sensor Data at a Nursing Home

In Japan, because of the declining birthrate and the increase in the elderly population, users of nursing care facilities are increasing, and there is a critical problem that the number of caregivers is insufficient. One way to solve this problem is to optimize the work for the staff in nursing care facilities. For this purpose, we are conducting research to sense the activities of the nursing staff. In doing so, the sensor data are collected using a smartphone or sensor device. Using the proposed method, we can expect that the data observations will become easier, and the data analysis more efficient.

We used SimpleLink SensorTag CC 2650 STK (Texas Instruments)[1] sensor devices, which were installed beside the beds in the individual rooms of the residents, or worn on the chests of the nursing staff. Smartphones with Android OS were installed in rooms where sensor tags were unavailable, and in shared locations. We used the illuminance sensor of the smartphone and the acceleration sensor of the sensor tag attached to the nursing staff.

At the same time, when the nursing staff carried out their work, they selected from about 25 action labels, including "patrol," "personal record (of their duties)," and "toilet assistance (of the residents)" from the smartphone app when they performed their activities. We extracted the acceleration sensor data from the time zones labeled "patrol" and "personal record." "Patrol" is a task of visiting the room of each resident and checking for abnormalities, which is an important activity to grasp the state of each resident. "Personal record" is a task to record such information as the body temperature and blood pressure of each resident, their physical condition, and other factors. It is also an important activity to improve the operational efficiency, such as how long it takes to record such information.

[1] http://www.ti.com/tool/cc2650stk.

4.2 Evaluation System

To conduct our experiments, we developed an evaluation system that operates on the Web, as shown in the Fig. 6. We used JavaScript as the applied language, namely, plotly.js[2] of the JavaScript library to visualize the data. In addition, Google Chrome was used as the applied browser. The center of the screen displays the visualized sensor data. The subject marks the attention points with a red frame for this visualized object. At this time, the coordinates of the red frame are also acquired. In addition, when a plurality of attention points exist, a red frame can be added by pressing "Add frame" button at the bottom of the screen, and marks can be applied as a focus points at a plurality of locations. When the subjects finish adding a red frame to all attention points, they can proceed to the next data by pressing "Next" button. The specific experiment method is described in the next section.

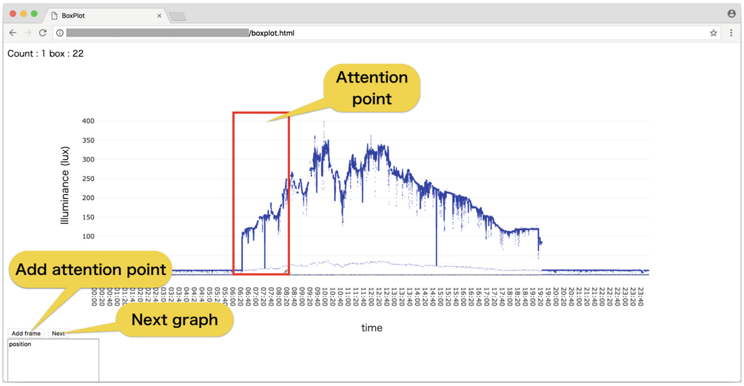


Fig. 6. Evaluation system (red frame, the attention points attached by the subject; "Add frame" button, addition of a red frame for the attention points; "Next" button, proceed to the next visualization data.) (Color figure online)

4.3 Evaluation Methods

The sensor data to be visualized are from the illuminance sensor and acceleration sensors and were obtained at the nursing home. We adopted these two types of sensor because the illuminance sensor data has a clear amount of changes, and the acceleration data has data of continuously fluctuate. The illuminance sensor data were visualized for one of 50 days of data, and the acceleration sensor data were randomly visualized from each of 20 data types labeled "patrol" and "personal record."

At that time, visualization was conducted using the two types of visualization methods considered for each sensor. For the illuminance sensor, as shown in

[2] https://plot.ly/javascript/.

Fig. 8, a box plot at 1 min intervals and a bar graph showing the average value at every 10 min were used, and for the acceleration sensor, as shown in Fig. 9, separate X-, Y-, Z-axis data and a three-axis composite value were visualized using a line graph. In the case of the illuminance sensor, by visualizing it with a box plot, it is possible to observe a change in data within a time zone. In addition, a bar graph is a commonly used visualization method in a wide range of fields [11].

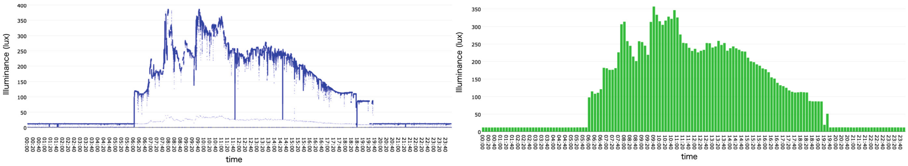
The subjects marked the attention points on this visualized graph with a red frame. At this time, the illuminance sensor divides the time width of a fixed length described in Sect. 3 (1) into each hour, and the acceleration sensor divides it into 2 min intervals.

Calculation of the feature vectors in Sect. 3 (2) is based on "Maximum (Max)", "Minimum (Min)", "Mean", "Median", "First quartile (1stQu)", "Third quartile (3rdQu)", "Variance (Var)", "Standard deviation (Sd)", "Number of data (Sample)".

Following Sect. 3 (3), the coordinate of the red frame is acquired, and T or F is determined. This task was carried out by the subjects (five male students, 23 in age), and we collected data on the attention points. Figure 7 shows the actual experimental scenery.

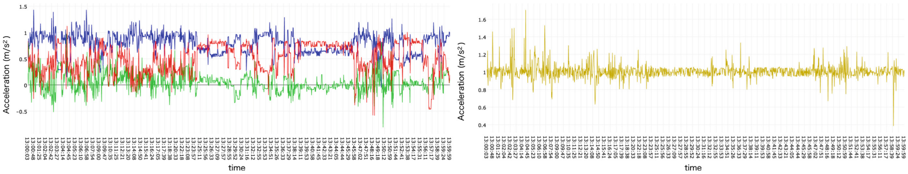
As a result, the illuminance sensor recorded 6,000 data, and the acceleration sensor recorded 1,300 data. We divided this created dataset into learning and test data, and we learned the data through a random forest method with T/F as the objective variables, and another feature quantity as the explanatory variables. At this time, the test data are one visualization data from one subject, and the learning data applies 1-user-image-leave-out cross validation for everything else. Specifically, we used the R language randomForest package in this paper. The evaluation item (1) was evaluated using this created algorithm.

Fig. 7. Landscape actually experimenting



Box plot Bar graph

Fig. 8. Example of visualization of illuminance sensor



Line graph (X, Y, Z) Line graph (three-axis composite value)

Fig. 9. Example of visualization of acceleration sensor

4.4 Results

Table 1 shows the results of item (1) with a cross validation for each visualization method of each sensor. When the illuminance sensor was visualized with a box plot, the accuracy was 85.7%, and when visualized with a bar graph the accuracy was 84.5%. When the acceleration sensor labeled patrol was visualized separately on the X, Y and Z axes, the accuracy was 85.0%, and when visualized with a three-axis composite value the accuracy was 87.9%. In addition, when visualizing the acceleration sensor data labeled as personal records separately for three axes, the precision was 76.6%, and the precision when visualized with a three-axis composite value was 80.8%.

Table 1. Accuracy for each visualization method. The "Sensor" column shows the sensor type, and the "Visualization Method" column shows the visualization method.

Sensor	Visualization method	Accuracy
Illuminance	Box plot	85.7%
	Bar graph	84.5%
Acceleration (patrol)	Line graph (X, Y, Z)	85.0%
	Line graph (three-axis composite value)	87.9%
Acceleration (personal record)	Line graph (X, Y, Z)	76.6%
	Line graph (three-axis composite value)	80.8%

Table 2 shows the result of item (2) with the estimation accuracy when applying the estimation model acquired by another visualization method. The first line shows the result of applying the model learned using the bar graph of the illuminance sensor to the box plot, where the estimation accuracy was 82.5%. The second line shows the opposite case of the first line, where the estimation accuracy was 81.8%. The third and fourth lines show the results of the acceleration sensor with "patrol" as the activity type. The third line shows the results of applying the model learned using the three-axis composite value as compared to those visualized separately on three axes, where the estimation accuracy is 70.9% and the estimation accuracy is higher than 50%, i.e., chance, for evaluation item (1), which is lower than the precision. The fourth row shows the opposite case, where the estimation accuracy is 71.6%, and the estimation accuracy is lower than the evaluation item (1). The fifth and sixth lines show the results of the acceleration sensor with "personal record" as the activity type. The estimation accuracy when applying the model learned using the three-axis composite value as compared to the three axes visualized separately was 68.8%. In the opposite case, the estimation accuracy was 69.9%, both of which are higher than chance; however, this is lower than the case of evaluation item (1). The reason for this lowering of the estimation accuracy will be discussed in the next section.

Table 2. Estimation accuracy when the estimation model of another visualization method is applied. The "Sensor" column indicates the sensor type. The "Training" column shows the visualization method for the training. The "Test" column shows the visualization method tested.

Sensor	Training	Test	Accuracy
Illuminance	Bar graph	Box plot	82.5%
	Box plot	Bar graph	81.8%
Acceleration (patrol)	Three-axis composite value	X,Y,Z	70.9%
	X,Y,Z	Three-axis composite value	71.6%
Acceleration (personal record)	Three-axis composite value	X,Y,Z	68.8%
	X,Y,Z	Three-axis composite value	69.9%

For the result of evaluation item (1), both the illuminance sensor and the acceleration sensor were able to obtain about an 80% estimation accuracy. For the result of evaluation item (2), with the illuminance sensor, the estimation accuracy did not decrease even when the learned model was replaced, whereas in the case of the acceleration sensor, the estimation accuracy was lower than that of evaluation item (1). The reasons for this decline in the estimation accuracy are discussed in the next section.

5 Discussion

In this paper, we applied two kinds of visualization using an illuminance sensor and an acceleration sensor, and evaluated the results using a random forest method applying the attention points as the object variable and the other feature vectors as the explanatory variable. With evaluation item (1), the illumination sensor and the acceleration sensor labeled "patrol" had an estimation accuracy of 80% or more. On the other hand, the acceleration sensor labeled "personal record" was slightly lower in estimation accuracy than the other two, which resulted from the user's attention points differing for each sensor.

Figures 10 and 11 show examples of which time zone the subject focused on in the data. With the illuminance sensor, it is understood that all subjects paid attention to places where the amount of change in data was large, as in Fig. 10.

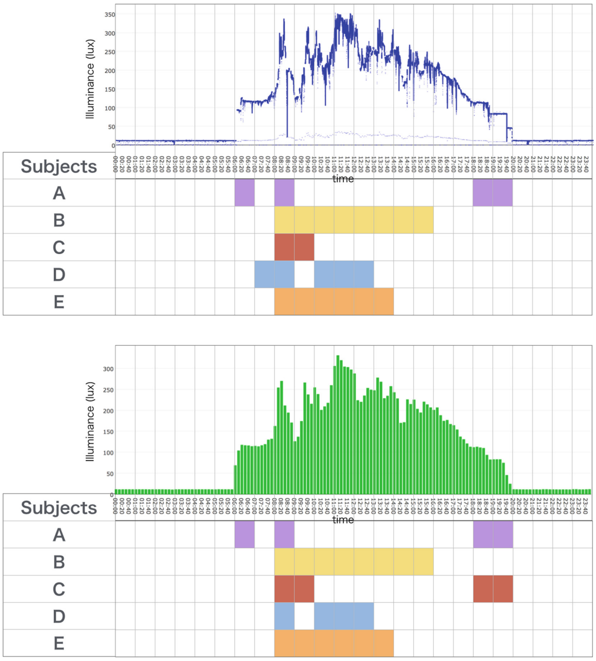


Fig. 10. Attention points for each subject (sensor type, illuminance sensor; upper half, visualization of the sensor data; lower half, attention points for each subject.)

Figures 12, 13 and 14 shows the importance of the feature vectors obtained when applying a random forest method. If we look at Fig. 12, we can see that the variance and standard deviation are particularly high. For the acceleration sensor labeled "patrol", we can see that all subjects paid attention to the same place, as shown in Fig. 11(Left). Therefore, it is considered that the estimation accuracy of the attention points increased. In the case of the acceleration sensor labeled "personal record", we can see that the attention points differ depending

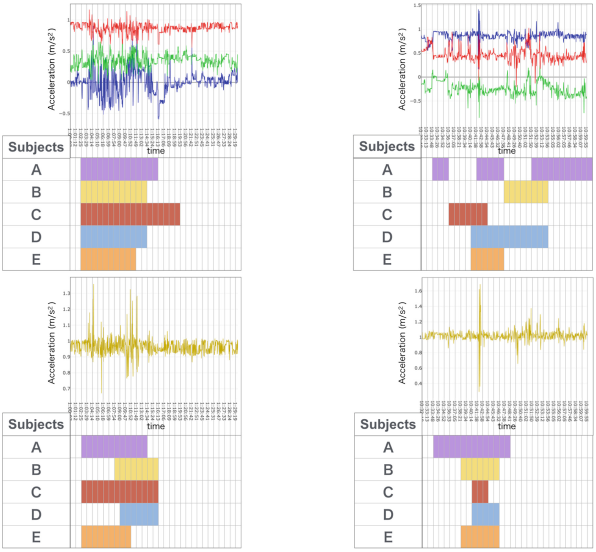
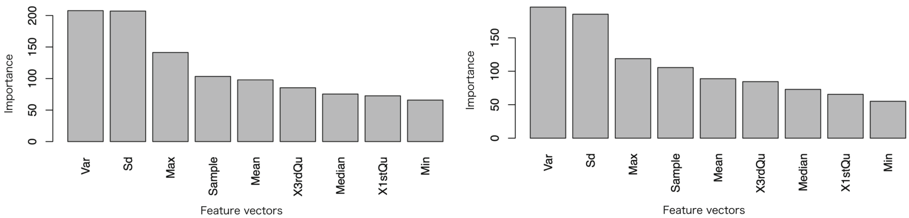


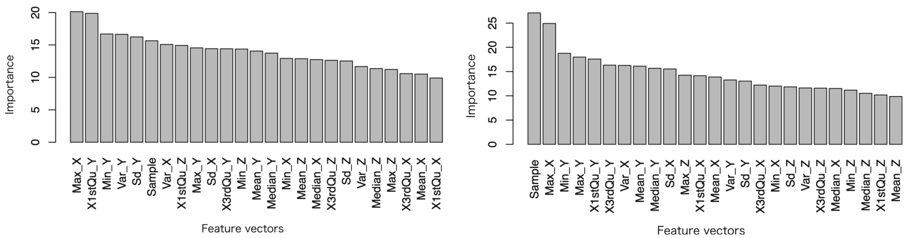
Fig. 11. Attention points for each subject (sensor type, acceleration sensor; left, patrol; right, personal record; upper half, visualization of the sensor data; lower half, attention points for each subject.)



Box plot

Bar graph

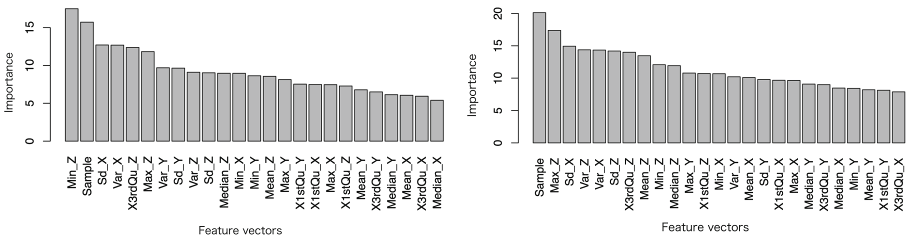
Fig. 12. Importance of feature vectors (illuminance)



Three-axes separately (X, Y, Z)

Three-axis composite value

Fig. 13. Importance of feature vectors in case of patrol (acceleration)

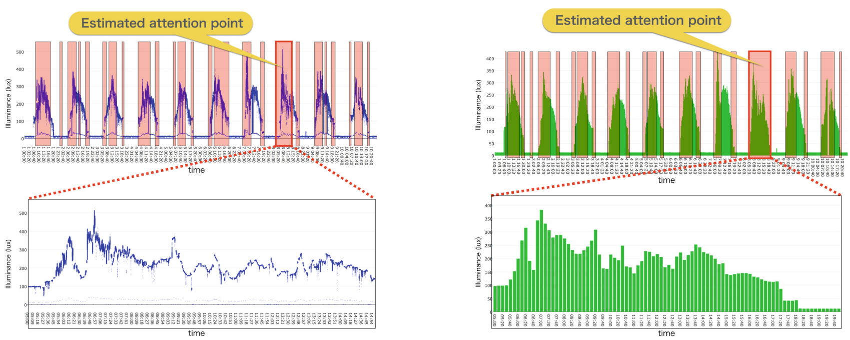


Three-axes separately (X, Y, Z) Three-axis composite value

Fig. 14. Importance of feature vectors in case of personal record (acceleration)

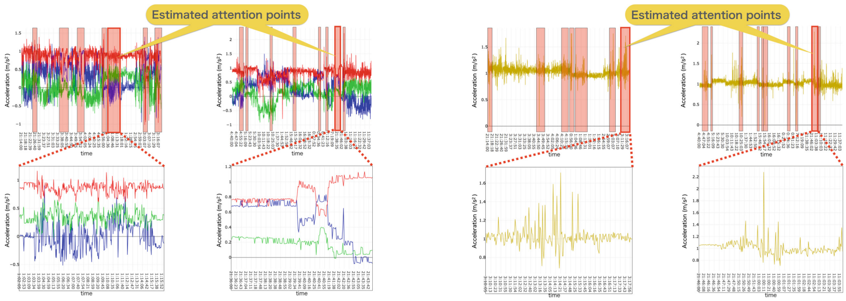
on the subject when the three axes are different, as shown in Fig. 11(upper right). Therefore, the estimation accuracy is considered to be less than 80%.

In addition, Figs. 15 and 16 shows a visualization of this estimation result. Figure 15 shows the result of visualizing the estimated portion of the illuminance sensor. It can be seen that the estimated range is wider than the attention points of the subject shown in Fig. 10. This is considered to be due to the conformity rates of 51.0% and 50.2% for the box plot and bar graph, respectively, and it is presumed that places other than the attention points were also estimated. Figure 16 shows the results of visualizing the estimated position of the acceleration sensor. In the case of the acceleration labeled "patrol", the precision of the three-axis composite value and the three separate axes are 75.8% and 78.3%, respectively, and most of the estimated points can be said to be attention points. The recall rates are 77.1% and 80.0%, and it is possible to estimate the majority of points that are actual attention points. In the case of acceleration labeled "personal record", the precision rates of the three-axis composite value and the three separately axes are 58.7% and 67.2%, and the positions where the



Box plot Bar graph

Fig. 15. Example of visualizing estimated results (illuminance: for the upper half, the part surrounded by the red frame was a place estimated as attention points; lower half, visualized by zooming in on the estimated attention points.) (Color figure online)



Three axes separately (X, Y, Z) Three-axis composite value

Fig. 16. Example of visualizing estimated results (acceleration: in the upper half, the part surrounded by the red frame was a place estimated as the attention points; the lower half, visualized by zooming in on the estimated attention points.) (Color figure online)

three separate axes do not have more attention points are estimated. However, the recall rates are 67.4% and 76.7%, which shows that we can estimate many points that are actual attention points.

For the illuminance sensor used in evaluation item (2), even when the learning algorithm was replaced, there was not much change in the estimation accuracy for evaluation item (1). This is almost the same for both visualization methods and attention points, as shown in Fig. 10. Furthermore, because the importance of the feature vectors of both figures is the same in Fig. 12, it is considered that the precision did not decrease even when adapting to another visualization method.

On the other hand, for the acceleration sensor labeled "patrol", the estimation accuracy was about 70%, which was lower than for evaluation item (1). In the case of three separate axes, as shown in Fig. 11(left), the range of the attention points are wide for each subject, whereas in the case of the three-axis composite value, the range of the focus area differs for each subject. In addition, although the feature vectors such as the maximum value of the X axis and the minimum value of the Y axis are both high, as indicated in Fig. 13, the importance of the other feature vectors are different, and was considered to have decreased.

In the case of the acceleration sensor labeled "personal record", the estimation accuracy was less than 70%. As we can see in Fig. 11, this shows that the attention points are different for each subject compared with "patrol". In addition, as shown in Fig. 13, the importance of the "number of samples" is high for both cases, although the other features are different. Therefore, when the learning algorithm is replaced, the estimation accuracy is considered to be lower than for evaluation item (1).

In this way, when the change in data, such as from the illuminance sensor, is clear, it is possible to estimate the attention points with high accuracy. Furthermore, when the visualization method is similar, it was confirmed that the method can be applied to multiple visualization methods using a single learned algorithm. It was also confirmed that it is possible to estimate the attention points with high accuracy even when the data constantly change, such as with an acceleration sensor.

6 Related Works

Various studies on visualization methods regarding points of viewer attention have been conducted. In the following, in Sect. 6.1, we describe research gathering attention points data and estimating the design optimization and attention points. Section 6.2 also describes various studies on visualization methods.

6.1 Visual Attention on Design

Research using data on viewer attention points includes an optimization of the design of a Web page by collecting human gaze data [4]. An excellent web design includes how information can be efficiently conveyed to people in a manner intended by the designer to achieve a certain purpose. Therefore, it is necessary to predict and design the areas of interest so that people can efficiently collect information. With this method, by designing a Web page as an input, it is possible to create a design that can easily guide people's attention while maintaining as much of the design as possible. However, it takes a significant amount of time to collect human gaze data, and Web designs given as input must be completed to a certain extent. Bylinskii et al. [5] proposed a method for estimating where people focus on a graphic design and data visualization, and express such estimations through a heat map. Instead of tracking the viewer's line of sight through data collection, we use a method that can obtain similar results as gaze tracking using a mouse click, called BubbleView [6]. This makes it possible to collect data in an efficient manner. However, because this method estimates the attention point and expresses the result as a heat map, we have not conducted a design optimization based on the estimation result. In our research, even if switching to another visualization method can cope without the collection of new attention point data, in these studies, it is necessary to gather attention point data for each visualization method.

6.2 Visualization Methods

Systems that recommend which visualization is appropriate when visualizing data have been proposed [7,8]. Multivariate data are given as an input, and a plurality of visualizations are automatically performed through the selection of a certain variable. This type of system recommends a type of visualization combined with other non-selected variables. However, in the case of time series data such as sensor data, in general, a line chart is often used. It is not meaningful to use this type of system because its visualization method is limited. Because this system has a limitation regarding the amount of data that can be given as an input, it is incompatible with a large variety and quantity of data. Other systems specialized for time series data have also been proposed [9]. A line chart is visualized by providing the time series data as an input. When observing this type of visualization data, by zooming in on the area of interest, only that part is visualized, and the part of the original visualization data that is zoomed is expressed in a tree. With this system, it is easy to grasp which part of the original data the zoomed area shows. However, when the amount of data becomes too great, it becomes difficult to grasp what is being drawn, and zooming becomes difficult. To solve such a problem, a method of visualizing time-series data in a three-dimensional space has been proposed [10]. Using this method, it is possible to visualize enormous amounts of sensor data, but when expressed in a three-dimensional space, it becomes difficult to observe when compared with the case on a two-dimensional plane. In these studies, automatically estimating and visualizing the attention points, as achieved in our research, is not possible.

7 Conclusion

In this research, we proposed and evaluated a method to automatically estimate attention points for sensor data using supervised machine learning. As a result, the estimation accuracy of the illuminance sensor and the acceleration sensor labeled "patrol" was about 85%, but in the case of the acceleration sensor labeled "personal record" it was slightly lower at about 80%. We also evaluated whether the algorithms learned using a single visualization method can be applied to other visualization methods. As a result, it was possible to obtain an estimation accuracy of 80% or more with the illuminance sensor even if the learning algorithm is switched. In the case of the acceleration sensor, a slightly decreased estimation accuracy of around 70% was achieved. In the case of the illuminance sensor, all subjects focused their attention on almost the same place, but in the case of the acceleration sensor, it was considered that the estimation accuracy was lowered because the attention location of the subjects was slightly different for each visualization.

In the future, it will be necessary to search for different visualization methods and feature vectors that can improve the estimation accuracy for acceleration data. We will experiment with other sensor data and visualization methods, and apply them using a single learning algorithm, aiming at automatic visualization according to the users.

References

1. Chahuara, P., Fleury, A., Portet, F., Vacher, M.: On-line human activity recognition from audio and home automation sensors: comparison of sequential and non-sequential models in realistic smart homes. J. Ambient Intell. Smart Environ. **8**(4), 399–422 (2016)

2. Ueda, K., Suwa, H., Arakawa, Y., Yasumoto, K.: Exploring accuracy-cost trade-off in in-home living activity recognition based on power consumptions and user positions. In: 14th IEEE International Conference on Ubiquitous Computing and Communications (IUCC 2015), pp. 1131–1137 (2015)
3. Bikakis, N., Sellis, T.: Exploration and visualization in the web of big linked data: a survey of the state of the art. In: LWDM (2016)
4. Pang, X., Cao, Y., Lau, R.W.H., Chan, A.B.: Directing user attention via visual flow on web designs. ACM Trans. Graph. (TOG) **36**, 240 (2016)
5. Bylinskii, Z., Kim, N.W., O'Donovan, P., Alsheikh, S., Madan, S., Pfister, H., Durand, F., Russell, B., Hertzmann, A.: Learning visual importance for graphic designs and data visualizations. In: Proceedings of 30th Annual ACM Symposium on User Interface Software & Technology (2017)
6. Kim, N.W., Bylinskii, Z., Borkin, M.A., Gajos, K.Z., Oliva, A., Durand, F., Pfister, H.: BubbleView: an interface for crowdsourcing image importance maps and tracking visual attention. ACM Trans. Comput.-Hum. Interact. **24**, 36 (2017). (A Special Issue)
7. Wongsuphasawat, K., Moritz, D., Anand, A., Mackinlay, J., Howe, B., Heer, J.: Voyager: exploratory analysis via faceted browsing of visualization recommendations. IEEE Trans. Visual. Comput. Graph. **22**, 649–658 (2015)
8. Wongsuphasawat, K., Qu, Z., Moritz, D., Chang, R., Ouk, F., Anand, A., Mackinlay, J., Howe, B., Heer, J.: Voyager 2: augmenting visual analysis with partial view specifications. In: Proceedings of 2017 CHI Conference on Human Factors in Computing Systems (2017)
9. Walker, J., Borgo, R., Jones, M.W.: TimeNotes: a study on effective chart visualization techniques for time-series data. IEEE Trans. Visual. Comput. Graph. **22**(1), 549–558 (2016)
10. Imoto, M., Itoh, T.: A 3D visualization technique for large scale time-varying data. In: 14th International Conference on Information Visualisation (IV10), pp. 17–22 (2010)
11. Borkin, M.A., Vo, A.A., Bylinskii, Z., Isola, P., Sunkavalli, S., Oliva, A., Pfister, H.: What makes a visualization memorable? IEEE Trans. Visual. Comput. Graph. **19**, 2306–2315 (2013). (Proceedings of InfoVis 2013)