Visible Position Estimation in Whole Wrist Circumference Device towards Forearm Pose-aware Display

Yuki Tanida¹, Kaori Fujinami²,∗

¹Department of Industrial Technology and Innovation, Tokyo University of Agriculture and Technology, 2-24-16 Naka-cho, Koganei, Tokyo 184-8588, Japan
²Department of Computer and Information Sciences, Tokyo University of Agriculture and Technology, 2-24-16 Naka-cho, Koganei, Tokyo 184-8588, Japan

Abstract

Smart watches allow instant access to information; however, the visual notification is not always reachable depending on the forearm posture. Flexible and curved display technologies can enable full-wrist circumference displays that show information at the most visible positions using pose awareness. A prototype device is implemented with 10 LEDs and 10 accelerometers around the wrist. The most visible LED is estimated using a machine learning technique. The main idea is to utilize direct relationship between the raw acceleration signals and the position of the most visible LED, rather than assigning the position by particular classes of activities or forward-kinematic model-based estimation. Also, sensor reduction is attempted by introducing new features. A user study showed that the system allowed 89.9% of the system’s judgment to fit with the gap of 1 LED (18 mm) from the user’s expectations. The rotation-sensitive features proved to be informative, and a single sensor placed on the inside of the wrist achieved a performance level on par (F-measure 0.681) with the performance when all (10) sensors are used.

Received on 23 June 2017; accepted on 27 September 2017; published on 14 March 2018

Keywords: wearable computing; flexible display; wrist-worn device; accelerometer; notification; LED; machine learning

Copyright © 2018 Yuki Tanida and Kaori Fujinami, licensed to EAI. This is an open access article distributed under the terms of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/), which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi:10.4108/eai.14-3-2018.154341

1. Introduction

Wearables attached to a user’s body as everyday objects, like a watch, a shirt, shoes, a bracelet, and so on, allow the user to be monitored and/or informed anytime, anywhere. Various such applications have been proposed for healthcare and fitness [11, 32, 45], social interaction [5, 20, 21, 25], work support [37], safety awareness [1, 15], education/training [8, 36] and notification of messages. In particular, wrist-worn devices are very popular, boosted by the cultural affinity to watches and bracelets. Wrist is considered to be the most noticeable and accessible position for information notification among other bodily places such as arm, chest, shoulder, thigh, waist, hip, and shoe [4, 16]. A number of such products that fit into the category of smart watches have been commercially brought to the market [3, 14, 34].

In a traditional smart watch that uses the watch face as a display area, displayed information can be hidden from the user’s viewpoint depending on the posture of the forearm, which causes a delay or missed information delivery. Tactile feedback, for example with vibrations, can draw the user’s attention without distracting the eye. However, these vibrations nonetheless distract the user from ongoing tasks because he/she needs to turn the watch display to his/her face to obtain information details. To overcome this limitation, flexible and curved displays have been drawing attention for using with wrist-worn devices [6, 30, 33, 39, 42]. Such devices expand the display area from a portion on the wrist, like watch face, to the whole wrist circumference. This motivated us to investigate a method to adjust the
placement of information display within the whole wrist circumference by taking into account the device posture. We believe that this approach will result in fewer interruptions to ongoing tasks.

With this goal in mind, we propose here a method to identify the most visible position on the wrist from among ten candidate positions based on the forearm posture. The main idea behind this approach is to utilize the relationship between the sensor (accelerometer) signals and the most visible position. Applications for such a forearm pose-aware information display are, for example, notifying the user of incoming message by turning the most visible part on (or changing its color), displaying the email icon with the number of unread messages, or displaying weather forecast. Our proposed method is designed to be used as an ambient or glanceable display [10, 18]. Detailed information containing longer text and complex figures should be explicitly accessed by the user after checking the initial information and deciding to suspend on-going task. Thus, orientation-adjustment is not considered because the aforementioned content is less likely to be affected by display orientation; this simplifies the system design. The contributions of the article are as follows:

- A LED-based full-wrist circumference display device is developed as a tool to investigate methods of identifying the most visible position based on the forearm posture.

- A method for selecting the most visible position based on machine learning is proposed. Using simple gravitational force as a classification feature achieves the result that 89.9 % of the system’s judgment fits within the width of ±1 LED (±18 mm) of the user’s expectations regarding the most visible LED.

- Classification features are explored to improve the selection accuracy as well as to reduce the number of sensors. The results show that the rotation-sensitive features are informative, and even a single sensor placed on the inside of the wrist achieves a performance level on par with the ten-sensor version.

The rest of this article is organized as follows. Related work is examined in Section 2. Section 3 describes a prototype system that determines the most visible position of a LED from ten candidate positions on the whole wrist circumference based on the forearm posture. An experiment to test the performance of the position selection is presented in Section 4. An attempt to reduce the number of sensors is described out in Section 5. Finally, Section 6 concludes the article with an outlook.

2. Related work

2.1. Interaction on the arm and hand

The wrist is visible in most hand activities. Harrison, et al. carried out an experiment to evaluate the reaction-time performance against visual stimuli at seven positions on the body: shoulder, chest, upper arm, waist, wrist, thigh, and top of the shoe [16]. The result showed that the wrist had the fastest reaction time to a visual stimuli with the smallest standard error compared to other positions. They argue that the wrist has a unique role to deliver information most saliently during hand activities. In another research, Ashbrook, et al. [4] carried out an experiment to investigate the effect of placements of on-body device (hip, pocket, and wrist) and activities (walking and standing) on response time, where audio notification (ringing) was utilized. The result showed that the wrist was the fastest in acknowledging the ringing. These studies suggest that the wrist is the most suitable place for quick access to information on the body.

In a conventional wristwatch-type device, the display area is limited and fixed, namely the watch face. Flexible and curved display devices [30, 33] expand display area from a portion on the wrist to whole circumference of the forearm [6, 39, 42]. Snaplet [39] is a sensor-augmented flexible display device that can detect the shape of the display, e.g., convex and flat, and offers different interactions depending on the shape. SleeD [42] is a touch-sensitive sleeve display that facilitates explicit interaction with a wall-type display. DisplaySkin [6] is a whole-wrist circumference display that is flexible [30], and aims at realize implicit interactions with the user by adjusting the position and the orientation of displayed information to the forearm posture relative to the shoulder. An experiment showed that the proposed pose-aware display reduced the time to acknowledge the notification, and minimized the overall interruption. The concept of the display position adjustment is shown in Cito [12], where the watch face itself can move around the wrist manually or automatically to provide a variety of new interactions. We share the idea of adjustment of display position with DisplaySkin, which comes from the nature of ambient or glanceable display; however, we limit the display to show only iconic images or a small number of characters not requiring an orientation adjustment. Also, to make the display less obtrusive, we do not take the watch-face movement approach like Cito, because moving a relatively large display with a certain mass would generate tactile feedback and cause distraction. This design choice makes our system simple yet useful as described in Section 2.2.

An array of tiny displays can create a flexible display environment [23, 28]. Facet [23] has six segments of tiny displays around the wrist; the user can assign particular applications, e.g., weather forecast, calendar, etc., to each of the segments. The range of the display area can...
be enlarged, and the position of the information can be switched depending on the frequency of use. However, the pairing of application (information) and the display position needs to be specified by the user manually. So, it may be difficult to acquire certain information depending on the posture. Our proposed method can be incorporated into Facet to make it pose-aware. AugmentedForearm [28] is realized by connecting four tiny displays in a row. In their prototype, information detail and the level of privacy is controlled depending on the coverage of clothes, e.g., sleeves, and the orientation to the user. However, a method of estimating the visibility of information to the user is not incorporated.

2.2. Human posture tracking

A number of human posture-tracking methods have been developed for human-computer interaction. A popular approach uses optical markers attached to the body and cameras installed in the environment to capture the posture [27, 41]. By contrast, the depth sensor-based approach is a markerless method [9, 43], in which the user does not need to attach any marker on his/her body. However, a camera, which is a depth sensor, still needs to be installed in the environment, which limits the utilization of the system in a specific area, and also faces occlusion problems.

Wearable devices have been investigated to allow a system to track the user’s body posture anytime, anywhere. Conductive polymer is used to track wrist motion such as pronation/supination, ulnar/radial deviation, and flexion/extension, in which load [22] and stretch [38] generate a change in resistance in response to the material deformation by wrist motion. Furthermore, stretch sensing based on fiber optics is applied to track flexion/extension and abduction/adduction of wrist [19]. The utilization of surface electromyogram (sEMG) sensor is a more direct approach [29, 40, 46], in which electrical currents generated during muscle contraction are measured, and neuromuscular activities that generate particular postures are captured through signal processing. By attaching widely spread electrodes, as in a whole-wrist circumference device, various wrist motions can be recognized [26, 29, 40]. These include flexion/extension and pronation/supination of wrist. The posture of forearm relative to the user’s eyes should be identified to adjust the position of the display to the users’ eye position. This requires separate devices for elbow and shoulder joint movements, which makes the system complicated and cumbersome for the user. Furthermore, resistive sensor-based and sEMG-based approaches require the sensor to be in contact with the wrist firmly to capture proper deformation of the resistive material and neuromuscular signal.

Inertial sensors provide information about the posture based on the measurement of accelerometer, gyroscope, and magnetometer. The advantage over resistive sensor-based and sEMG-based approaches is that it allows contact-free posture measurement. This means that an inertial sensor does not need to be in contact with the skin, but can be attached on the body part of interest. By attaching two devices, for example one on the upper arm and the other on the forearm, and setting body parameters such as the relative position of the eyes from the shoulder and the length of both upper arm and forearm, the forward kinematic model can infer the position and the orientation of the wrist relative to the shoulder. This allows a system to identify the most visible position as well as the orientation in any position [6]. For our research, we adapt the inertial sensor-based approach. However, we use only one device that consists of one or more accelerometers on the wrist, as this can fulfill the requirement for our applications: the display content in our system does not contain long pieces of text, but iconic images or a few words. So, a user may interpret the information without the system having to adjust the orientation relative to the user’s eyes. In our approach, the model of visible positions for different forearm postures is constructed by learning the relationship between the sensor signals that represent particular postures and the most visible positions. Such “evidence-based” model requires a wide variation in pairing of sensor signals and the most visible positions; however, the processing algorithm becomes very simple, and the system’s output is more effective because it reflects actual visibility from the user’s point of view.

3. Prototype System

A prototype system is implemented based on our ideas to investigate the most-visible-position-detection method.

3.1. Design

Assumptions and Requirements. The design of our prototype system is based on the requirement that to estimate the forearm posture, the user would wear a single device on his/her wrist, rather than attaching multiple devices (as in [6]). The device is assumed to be stable around the forearm, meaning that it does not rotate significantly around the wrist during forearm motion, and therefore the display position on the device does not change. This assumption is also followed in the wrist watch design. We further require that the device be flexible like a wrist belt. For judging the visibility of the display, we assume that the forearm posture will be stable for a certain time period. This assumption is also often made in commercial smartphones [2, 13] on the basis of mobile human-computer interaction research [17, 35], where the screen orientation, namely landscape and portrait, changes depending on the orientation of the device.
Design Choices. One sensor may be attached anywhere to represent the “state” of the device, and to find the relationship between the state and visibility if the device is rigid and is circular or elliptical. Here, the state of the device is defined by the shape and the posture of the device. As assumed above, the device is flexible. Therefore the shape of the device may change during wearing and may differ from individual to individual. To address the issue, two solutions are possible: 1) finding places where deformation hardly happens and 2) capturing the dynamic shape changes. We can make an analogy with sEMG-based human motion tracking, in which specific electrode positions are utilized [7], or dense electrodes on a whole-wrist circumference device are used to measure forearm activity [29]. We took the latter approach to capture the shape of the device and the posture using as many accelerometers as possible. This means that the shape of the whole-wrist circumference device is represented by a set of vectors of three gravitational components from all accelerometers. However, using many sensors is not the final choice; the number of electronic components should be minimized to reduce the cost and the power consumption and also to minimize the size and the weight of the device. So, following an approach similar to that of the whole wrist circumference sEMG-sensing device [29], we also tried to reduce the number of sensors, as discussed later in Section 5.

The final design of the display would be realized by a single continuous wrist-circumference display [30, 33], on which icons of corresponding data sources will be presented with additional information at the most visible position. However, in our prototype, multiple LEDs are aligned around the wrist side-by-side at equal intervals. The LED positions in this prototype can be converted to particular areas on the continuous display concatenated with each other.

The components of the prototype system are illustrated in Figure 1. The system identifies the most suitable LED for a particular forearm state based on the relationship with acceleration signals. The main idea is that the choice is made based on specific patterns of acceleration signals, rather than specific class of activities like reading a book or driving a car. The system does not need to recognize the type of activity and obtain the most suitable position for the activity. This frees the device manufacturer from the burdensome task of having to collect data for each activity class. The process of obtaining one particular position (ID) of LED from among multiple LEDs is defined as a classification task in supervised machine learning, in which a set of data (classification features) are given to a trained classifier, which outputs the ID of the chosen LED.

3.2. Implementation

A prototype system was implemented as follows. Ten full-color LEDs (SWITCH SCIENCE SSCI-019156) and ten tri-axis accelerometers (ADXL335) are connected to one Arduino UNO with Mux Shield II via a soft cable of about 1.2 m, in which accelerometers are sampled at a rate of 50 Hz. Mux Shield II is utilized to increase the analog input ports. Arduino is connected to a PC via USB cables. It runs the feature calculator and the most-visible-LED-identification component (hereinafter we call the identification component “LED selector”). In the data collection and experiment, Arduino is usually put on a table and, if necessary, the user holds it in the other side of the hand. In this prototype, mean values of ten samples (=200 ms) for each of the three axes of the ten accelerometer are utilized as features that represent gravitational force for each axis. We consider that the combination of component gravitational forces of three axes represent the posture of that accelerometer sensor, and that the postures of the ten sensors around the wrist indicate the forearm posture in the global coordinate system. So, the dimension of the feature vector is 30 (=10 sensors × 3 axes). A Weka machine learning toolkit (version 3.7.13) is used to model and test the LED selector.

Figure 2 depicts the equipment of the system showing the ID assignments for both accelerometers and LEDs. The sensors and the LEDs are soldered on a flexible circuit board, and a polyurethane sheet (5 mm thickness) is placed between the circuit board and the skin to avoid hurting the skin in case that the board directly touches the wrist. So, the thickness of the device from the skin to the top of the accelerometer is about 13 mm. Due to the size limitation of the sensor and LED modules, the accelerometers and LEDs are placed at average intervals of 18 mm and 17 mm, respectively. As the device is assumed to be stable around the forearm, the IDs correspond to the positions of accelerometers and LEDs: the ID 1 is on the ulnar styloid process (marked “1” in Figure 2), and the numbers increase in a clockwise manner. This means that LED “6” is near the radius styloid process, while LED “9” is around the inner-wrist center. Note that there is a 1 mm difference in the size of the sensor and the LED, which breaks a one-to-one
relationship between these components. However, this is not problematic because these components are placed independently placed: that is, sensor$i$ is not intended to indicate the posture of LED$i$, but to estimate the posture of forearm.

![Figure 2. Equipment of the prototype system and the assignment of the IDs of sensors and LEDs.](image)

3.3. Implementing LED selector

**Data collection.** As described above, the LED selector is realized by supervised machine learning, in which a classification model is trained with feature-class pairs. We collected a dataset of 18 postures in daily activities from 20 persons (16 males and 4 females). Their wrist circumference ranged from 13.0 cm to 17.5 cm (mean: 15.4 cm, S.D ±1.31). The device was attached on the non-dominant hand (the left hand for all the participants) to simulate a wrist watch. The participants were asked to perform each of the 18 postures five times (from posture 1 to 18) in their own way for one second per posture. The postures as seen in daily activities are shown in Figure 3. These are categorized into four groups considering the height of the forearm against the eyes as follows: 1) on or slightly above the thigh, 2) on or slightly above the table, 3) above the table (near the chest), and 4) higher than the eye level. In 4), we ask the participants to grab an object that is easily reachable, but not to try to grab too high.

The procedure for data collection was as follows. First, the participant was asked to simulate the activity and hold the posture with his/her forearm for one second. Then an experimenter turned on all LEDs. To identify a particular LED easily, each of the ten LEDs had a different color: green, orange, purple, red, blue green, blue, yellow, pink, light blue, and white. Once the most visible LED was answered, all LEDs were turned off, and another trial started from the beginning after an interval of 6 seconds. The participants were asked not to focus on the device until it turned on, but to perform the directed activity naturally. This procedure allows us to collect pairs of acceleration signals and class labels for supervised learning. Note that the LEDs were assigned colors in a random order for each trial, and we confirmed that the colors were distinguishable from each other. The participant could take rest between trials to avoid fatigue. No participant reported fatigue during or at the end of the experiment.

We performed a one-way ANOVA to check if there is any difference in the most visible positions among postures. A significant difference was found (F(17, 1782)=74.77, p<0.01). So, we conclude that the effectiveness of notification can be improved by displaying it at the most visible position for each posture. Note that we are interested in associating the acceleration signals (rather than a posture like reading a book, typing on a keyboard) with the position of the most visible LED. This has the advantage that the LED selector model can be applied to an unknown activity with similar postures.

**Classifier selection.** There are benefits and limitations associated with different classification models [31]. We compared various types of modern supervised classification models to find the most accurate one for later experiments. Naive Bayes (NB, a baseline approach), Nearest Neighbor (NN, instance-based approach), Multi-Layer Perceptron (MLP, an artificial neural network approach), Sequential Minimal Optimization (SMO, a support vector machine approach) with polykernel, decision tree (J48) and Random Forest (an ensemble learning approach) were used in the comparison.

To evaluate our method under realistic conditions, we conducted the leave-one-person-out cross-validation (LOPO-CV), where the data of one person is removed and the data of remaining persons are used to train the classifier. The data of the removed person is used to test the classifier. Then we start with the original dataset again (putting the data of the removed person back in the dataset), and remove the data of another person and repeat the training and the testing process. This is repeated until the data of each person has been removed and used for testing. In the end, the average of all these test is calculated. An initial analysis of the data showed that there was no preference for LED 1 and 10, which were located around ulnar styloid process, i.e., on the little finger side. So, the selector was trained for the remaining eight classes.

The main metric in the evaluation is the F-measure that is the harmonic mean between recall and precision. The metrics for each class are represented by Formulas (1) to (3), where $N^\text{correct}_i$, $N^\text{estimated}_i$, and $N^\text{judged}_i$ represent the number of cases correctly classified into class$_i$, the number of test cases in class$_i$, and the number of cases classified into class$_i$, respectively, while $i$ ranges...
over each of the eight classes.

\[
\text{recall}_i = \frac{N_{\text{correct}, i}}{N_{\text{tested}, i}} \quad (1)
\]

\[
\text{precision}_i = \frac{N_{\text{correct}, i}}{N_{\text{judged}, i}} \quad (2)
\]

\[
\text{F-measure}_i = \frac{2}{\frac{1}{\text{recall}_i} + \frac{1}{\text{precision}_i}} \quad (3)
\]

Table 1 summarizes the averages of recall, precision, and F-measure and relevant classifier parameters used in Weka. We applied grid search to find parameters for SMO. The table shows that SMO had the highest values in Recall, Precision, and F-measure; RandomForest was the second highest in Recall and F-measure; and Naïve Bayes was the second highest in Precision. Paired t-tests between the highest classifier (SMO) and these second highest classifiers showed that SMO is significantly higher than the second highest classifiers in F-measure (t(19)=4.53, p<0.01) and Recall (t(19)=5.43, p<0.01), but not significantly higher in Precision (t(19)=1.39, p>0.05). Given that the F-measure represents the overall performance of a classifier, we utilize SMO in the later experiment.

4. Experiment

After implementing the LED selector, we conducted an experiment to confirm the effectiveness of machine learning-based selection of the most visible LED. Also, the robustness of selection against various forearm postures was evaluated.

4.1. Method

In Section 3.3, an SMO-based LED selector selector, was built using the dataset from 19 users (after removing the data of one user (P), whose data was utilized for...
Table 1. Comparison of various supervised classifier approaches using Leave-One-Person-Out cross-validation (LOPO-CV).

<table>
<thead>
<tr>
<th>Method</th>
<th>Recall</th>
<th>Precision</th>
<th>F-measure</th>
<th>Relevant parameters in Weka</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naïve Bayes</td>
<td>0.619</td>
<td>0.667</td>
<td>0.630</td>
<td></td>
</tr>
<tr>
<td>NN</td>
<td>0.563</td>
<td>0.592</td>
<td>0.562</td>
<td>number of neighbors=1,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>nearest neighbor search=linier search</td>
</tr>
<tr>
<td>MLP</td>
<td>0.607</td>
<td>0.640</td>
<td>0.600</td>
<td>learning rate=0.3, epoch=500, momentum=0.2</td>
</tr>
<tr>
<td>SMO</td>
<td><strong>0.688</strong></td>
<td><strong>0.692</strong></td>
<td><strong>0.676</strong></td>
<td>kernel=poly kernel, complexity=0.05, tolerance=0.005, epsilon=1.0×10^{−12}</td>
</tr>
<tr>
<td>J48</td>
<td>0.572</td>
<td>0.598</td>
<td>0.570</td>
<td>confidence=0.25, minimum number of instances per leaf=2</td>
</tr>
<tr>
<td>RandomForest</td>
<td>0.640</td>
<td>0.659</td>
<td>0.634</td>
<td>number of trees=50</td>
</tr>
</tbody>
</table>

Similar to as described in Section 3.3, the device was attached to the users on their non-dominant wrists (left hands), and the users were asked to perform 18 postures (see Figure 3), five times each, in their own ways. For each trial, the system determined one LED as the most visible one and turned it on in red color. The users were asked to answer the level of visibility based on the criteria shown in Table 2. Although the device consisted of 10 LEDs, we asked the users to assume that they were wearing a display showing textual and graphical information with small size words or images, and they can imagine the contents, the font size, background/foreground color, and the message length of as they wish. So, the difference between “2” and “3” ratings is whether the users consider the text presented by their assumed applications is readable (“3” rating) or not (“2” rating). To obtain the ground truth, the users were asked the most visible positions before the system showed its decision. This is regarded as the user “expectation”.

4.2. Results

Performance of visible LED selection. The percentage figures in Table 3 indicate the agreement of the user preference against the system’s judgment. The highest percentage of “4” rating, which indicates the best visible position for the users, is 48.0 %, which appears in postures 6 and 13, while the lowest percentage (14.0 %) is found in posture 17. We also show the performance under a relaxed condition, where “3” and “4” ratings are merged into a single category “adequate position”. For this category, the highest percentage (100 %) appears in postures 6, 13 and 15, while the lowest percentage (79.0 %) is found for postures 17 and 18 . As shown in the table, the overall average of best position selection is 35.9 %, while that of adequate position is 94.9 %.

One-way ANOVA for the best position selection shows a significant difference among posture groups (F(3, 356)=7.55, p<.01). Regarding the adequate position selection, one-way ANOVA also shows a significant difference among posture groups (F(3, 356)=141.23, p<.01). Finally, a significant difference among posture groups is found in the average ratings (the rightmost
column) as a result of one-way ANOVA (F(3, 356)=16.95, p<.01).

We examined the correlation between the wrist circumference and the system’s judgment. Pearson’s correlation coefficients between the wrist circumference and 1) the ratio of best position selection, 2) the ratio of adequate position selection, and 3) average rating are shown in Table 4. The table indicates that the percentages of best position selection and adequate position selection decrease as the wrist circumference gets longer, and that the average rating also decreases. However, t-tests for correlation revealed that they are not significantly correlated (p>.05).

The gap between the system’s judgment and the user expectations. We show the gap between the system’s judgments and the user expectations, which is defined by Formula (4) as the subtraction of the LED ID that the system selected as the best one (LED_{system}) from the LED ID that the user considered the best (LED_{subject}). Although the values of LED_{user} and LED_{system} are primarily nominal, they can be regarded as the distance from the starting point, i.e., ulnar styloid process, because the LEDs are arranged at equal interval. As described in Section 4.1, users were asked to select the most visible position (LED ID) prior to the system’s judgment. So, a positive gap value indicates that the system selected a LED more towards the pinkie (little finger) than the user expectation when the palm was down. Zero means that the system’s response and the user expectation agreed for the most visible LED position.

\[ gap = LED_{user} - LED_{system} \] (4)

Table 5 shows the detail of the relative frequency of gaps (percentages) per posture. An average gap is defined as the average of gaps among all users. The difference between Tables 3 and 5 is that Table 3 indicates the user’s agreement with the system’s judgment, while Table 5 shows the approximate physical distance between the system’s judgments and the user expectations in terms of LED IDs. Given that the average gaps are positive, we can see that the system tended to select on the left of the position expected by the user. One-way ANOVA regarding average gap shows a significant difference among posture groups (F(3, 356)=7.95, p<.01).

### 4.3. Analysis

In the user experiment, we utilized the LED selector built in Section 3.3. So, it has an average recall, precision, and F-measure of 0.688, 0.692, and 0.676, respectively. However, taking into account the fact that there are practically eight classes, the average “best position” performance of 35.9 % in Table 3 is much higher than random selection (=12.5 % (1/8)). Also, the average performance of 94.9 % for “adequate position” indicates that the system works well to select the best display location according to the user expectation.

In Table 5, 35.9 % of the system’s judgments has zero gap, and 89.9 % of them fit in the gap of ±1 LED on overall average. This corresponds to a range of 36 mm (the distance between three LEDs). An average gap of 0.73 especially indicates that the system tends to turn on a LED that is placed outside the user expectation within one LED (13.14 mm).

As shown in Sections 4.2, there is some variation in the performance among different postures. Post-hoc test regarding the difference of the best position selection (rating 4) by posture groups shows that Group 1 is significantly lower than Groups 2 and 3 (p<.01), and Group 4 is significantly lower than Groups 2 and 3 (p<.01). In the performance of adequate position selection (ratings 3 and 4), Group 4 is significantly lower than the other three groups (p<.01), and Group 2 is significantly lower than Group 3. In the average ratings, Group 4 received a significantly smaller average rating than the other three groups (p<.01), and Group 1 is lower than Group 3. Finally, Group 4 had a significantly larger gap than the other three groups (p<.01).

The groups were formed based on the height of the forearm against the eyes: for Group 1 this is the lowest while for Group 4 this is the highest. Based on the post-hoc tests, we found that the system works better for middle height (on the table, near the chest as for Groups 2 and 3, especially for Group 3) than the lowest height (on or slightly above the thigh, as for Group 1, or higher than eye level, as for Group 4). The LED selector was built using the dataset from all other users except for the user who was to be tested. So, we consider that the low performance was mainly caused by the failure in handling individual differences. Middle height indicates that the distance between the device and the user’s eyes is small; we assume that the visibility might not vary much among the users. In addition to the distance problem, we consider that the lowest performance in Group 4 might be caused by an insufficient number of training instances. As shown in Figure 3, postures 17 and 18 in Group 4 look different from each other as well as from the other postures. By contrast, postures 1, 3, and 4 look very similar to each other, which triples the size of the dataset for these postures and allows the LED selector to weaken individual differences. To confirm these hypotheses, more controlled experiments are needed.

As described in Section 4.1, the difference between ratings 2 and 3 is whether the user considers an imaginary text readable or not. We had a follow-up interview by e-mail, and found that the users mostly assumed a text of 5 Japanese and Chinese characters (a new line is started for more than 5 characters) with
Table 3. The percentage of user judgments that fit into “best position” and “adequate position” [%]. Also, average ratings are shown.

<table>
<thead>
<tr>
<th>Group</th>
<th>Posture</th>
<th>Rating 4 only (Best position)</th>
<th>Ratings 3 and 4 (Adequate position)</th>
<th>Average rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>20.0</td>
<td>97.0</td>
<td>3.2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>31.1</td>
<td>96.0</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>39.0</td>
<td>99.0</td>
<td>3.4</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>34.0</td>
<td>98.0</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>27.0</td>
<td>99.0</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td>average</td>
<td>30.2</td>
<td>97.6</td>
<td>3.3</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>48.0</td>
<td>100.0</td>
<td>3.5</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>43.0</td>
<td>96.0</td>
<td>3.4</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>31.0</td>
<td>89.0</td>
<td>3.2</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>37.0</td>
<td>98.0</td>
<td>3.4</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>45.0</td>
<td>88.0</td>
<td>3.3</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>35.0</td>
<td>97.0</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td>average</td>
<td>39.8</td>
<td>94.6</td>
<td>3.3</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>33.0</td>
<td>97.0</td>
<td>3.3</td>
</tr>
<tr>
<td>13</td>
<td>13</td>
<td>48.0</td>
<td>100.0</td>
<td>3.5</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>45.0</td>
<td>98.0</td>
<td>3.4</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>45.0</td>
<td>100.0</td>
<td>3.5</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>39.0</td>
<td>99.0</td>
<td>3.4</td>
</tr>
<tr>
<td></td>
<td>average</td>
<td>42.0</td>
<td>98.8</td>
<td>3.4</td>
</tr>
<tr>
<td>17</td>
<td>17</td>
<td>14.0</td>
<td>79.0</td>
<td>2.9</td>
</tr>
<tr>
<td>18</td>
<td>18</td>
<td>32.0</td>
<td>79.0</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>average</td>
<td>23.0</td>
<td>79.0</td>
<td>3.0</td>
</tr>
<tr>
<td>Overall</td>
<td>average</td>
<td>35.9</td>
<td>94.9</td>
<td>3.3</td>
</tr>
</tbody>
</table>

Table 4. The relationship between the wrist circumference and the system’s performance

<table>
<thead>
<tr>
<th>Rating 4 only (Best position)</th>
<th>Ratings 3 and 4 (Adequate position)</th>
<th>Average rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.300</td>
<td>-0.206</td>
<td>-0.251</td>
</tr>
</tbody>
</table>

5.1 mm square each character on average. They expected the display to show the name or the type of application, e.g., Twitter, LINE, e-mail, news, as well as the sender of the message, a news headline, or the subject of an e-mail message. Five characters corresponds to 25.5 mm, which is equivalent to two to three LEDs, considering that a LED is 18 mm. So, we consider that the result obtained above represents a strict case for presenting an iconic image for notification with the size of a LED because of lower possibility of occlusion.

5. Reducing the number of sensors with new features

As described in Section 3.1, ten accelerometers were utilized to understand the posture, resulting in a dense sensor environment. However, for a practical system, the number of sensors should be as small as possible, while keeping the performance of the suitable-LED-selection system high. In this section, we attempt to reduce the number of sensors by introducing new features in addition to the original raw gravitational signals, so that each sensor contributes more information to the system.

5.1. Defining new features

We define new features by taking into account relationships between the axes and the dominance of particular axes for particular postures. Nine types of features are calculated from raw acceleration signals \( (raw_{x,y,z}) \), in which \( i \) (from 1 to 10) denotes the number of accelerometer, as shown in Figure 2. Formulas (5) to (13) define these features. Mean and variance represent the average and the variance of the three axes raw data, which are defined by the formulas (5) and (6), respectively. Gravitational acceleration is considered to be the raw data. The sign of the signal depends on the direction of the accelerometer toward the force of gravity although the magnitude of the force may be identical. So, we consider that the mean represents the general sensor direction, not the exact direction of each axis. The variance shows how the three axes differ from each other, which may contribute to the general sensor direction as well. Max (Formula (7))
Table 5. The detail of the gap between the system’s judgments and the user expectations per posture [%].

| Group | Posture | Average gap | < -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |<-
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>---</td>
</tr>
<tr>
<td>1</td>
<td>0.81</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>20.0</td>
<td>79.0</td>
<td>1.0</td>
<td>0.0</td>
<td>0.0</td>
<td>---</td>
</tr>
<tr>
<td>2</td>
<td>0.73</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>31.0</td>
<td>65.0</td>
<td>4.0</td>
<td>0.0</td>
<td>0.0</td>
<td>---</td>
</tr>
<tr>
<td>3</td>
<td>0.53</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.0</td>
<td>39.0</td>
<td>54.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.70</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>34.0</td>
<td>62.0</td>
<td>4.0</td>
<td>0.0</td>
<td>0.0</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.80</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.0</td>
<td>27.0</td>
<td>63.0</td>
<td>9.0</td>
<td>0.0</td>
<td>0.0</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>average</td>
<td>0.71</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.4</td>
<td>30.2</td>
<td>64.6</td>
<td>3.6</td>
<td>0.0</td>
<td>0.0</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.53</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>43.0</td>
<td>51.0</td>
<td>6.0</td>
<td>0.0</td>
<td>0.0</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.73</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>32.0</td>
<td>48.0</td>
<td>4.0</td>
<td>0.0</td>
<td>0.0</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.53</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.0</td>
<td>39.0</td>
<td>54.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.63</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>45.0</td>
<td>39.0</td>
<td>14.0</td>
<td>0.0</td>
<td>0.0</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.73</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>35.0</td>
<td>58.0</td>
<td>6.0</td>
<td>0.0</td>
<td>0.0</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>average</td>
<td>0.70</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.3</td>
<td>40.0</td>
<td>49.5</td>
<td>9.7</td>
<td>0.5</td>
<td>0.0</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.76</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>33.0</td>
<td>58.0</td>
<td>9.7</td>
<td>0.5</td>
<td>0.0</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.54</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.0</td>
<td>48.0</td>
<td>47.0</td>
<td>4.0</td>
<td>0.0</td>
<td>0.0</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.76</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>45.0</td>
<td>49.0</td>
<td>6.0</td>
<td>0.0</td>
<td>0.0</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.58</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.0</td>
<td>45.0</td>
<td>49.0</td>
<td>6.0</td>
<td>0.0</td>
<td>0.0</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.61</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>45.0</td>
<td>49.0</td>
<td>6.0</td>
<td>0.0</td>
<td>0.0</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0.69</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>39.0</td>
<td>53.0</td>
<td>8.0</td>
<td>0.0</td>
<td>0.0</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>average</td>
<td>0.64</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.4</td>
<td>42.0</td>
<td>51.2</td>
<td>6.4</td>
<td>0.0</td>
<td>0.0</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0.79</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>14.0</td>
<td>48.0</td>
<td>31.0</td>
<td>5.0</td>
<td>3.0</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0.79</td>
<td>1.00</td>
<td>2.0</td>
<td>8.0</td>
<td>32.0</td>
<td>31.0</td>
<td>18.0</td>
<td>5.0</td>
<td>3.0</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>average</td>
<td>1.01</td>
<td>0.50</td>
<td>1.0</td>
<td>5.0</td>
<td>23.0</td>
<td>39.5</td>
<td>24.5</td>
<td>5.0</td>
<td>1.5</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>0.73</td>
<td>0.10</td>
<td>0.1</td>
<td>0.9</td>
<td>35.9</td>
<td>53.1</td>
<td>8.7</td>
<td>0.7</td>
<td>0.2</td>
<td>---</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

and min (Formula (8)) provide the maximum value and minimum value, respectively, while “arg max” (Formula (9)) and “arg min” (Formula (10)) give information of the max and min axes, respectively. We consider that the pairs of max(min) and “arg max(min)” represent the information of dominant axis as a characteristic of the sensor posture. Finally, three types of ratios are defined as indicators of the relationships between the two axes (Formulas (11) to (13)).

5.2. Offline experiment

A LOPO-CV was carried out using the same dataset and SMO classifier as in Section 3.3. Table 6 summarizes the average F-measures of 20 users for each sensor. Additionally, in the rightmost column, the F-measures of the LED selector built from all the ten sensors is shown for comparison. As shown in the table, Sensor 7 had the highest F-measure (among the ten sensors) of 0.681, which is at the same level as all the sensors (“All”). Paired t-tests between “All” and individual sensors show that significant differences below p<0.05 are found in Sensors 1, 2, 3, 5, and 6. Furthermore, we could not find any significant difference among sensors 4, 7, 8, 9, and 10 (F(4, 95)=0.24, p>.05). As shown in Figure 2, and based on our observation of users with various wrist circumferences (from 13.0 to 17.5 cm), sensors 7 to 10 are located on the inside of the wrist. So, we can conclude that the system can be realized with only one sensor attached on the inside of the wrist.

\[ \text{mean}_i = \frac{1}{3} \sum_{a \in \{x,y,z\}} \text{raw}_{i,a} \]  
\[ \text{variance}_i = \frac{1}{3} \sum_{a \in \{x,y,z\}} (\text{raw}_{i,a} - \text{mean}_i)^2 \]  
\[ \text{max}_i = \max(\text{raw}_{i,a}, a = x, y, z) \]  
\[ \text{min}_i = \min(\text{raw}_{i,a}, a = x, y, z) \]  
\[ \text{max-axis}_i = \arg \max_{a \in \{x,y,z\}} \text{raw}_{i,a} \]  
\[ \text{min-axis}_i = \arg \min_{a \in \{x,y,z\}} \text{raw}_{i,a} \]  
\[ \text{ratio}_{i,x/y} = \frac{\text{raw}_{i,x}}{\text{raw}_{i,y}} \]  
\[ \text{ratio}_{i,x/z} = \frac{\text{raw}_{i,x}}{\text{raw}_{i,z}} \]  
\[ \text{ratio}_{i,y/z} = \frac{\text{raw}_{i,y}}{\text{raw}_{i,z}} \]  

We choose sensor 7, which has the highest F-measure in this sensor group, as a representative of effective sensors. An average F-measure using only original 3-axes acceleration signals of Sensor 7 is 0.659, which is significantly lower than that of the all-sensors version.
Table 6. Average F-measures of 20 users for individual sensor and all sensors (rightmost column). A single asterisk and a double-asterisk indicate that there is significant difference between the F-measures of “All” and a particular sensor at the confidence level of p<.05 and p<.01, respectively.

<table>
<thead>
<tr>
<th>Sensor ID</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-measure</td>
<td>0.631*</td>
<td>0.617*</td>
<td>0.623**</td>
<td>0.663</td>
<td>0.628**</td>
<td>0.659*</td>
<td>0.681</td>
<td>0.667</td>
<td>0.662</td>
<td>0.650</td>
<td>0.681</td>
</tr>
</tbody>
</table>

Table 7. Information gain of each feature obtained from Sensor 7 [bit]

<table>
<thead>
<tr>
<th>raw_x</th>
<th>raw_y</th>
<th>raw_z</th>
<th>mean</th>
<th>variance</th>
<th>max</th>
<th>min</th>
<th>max_axis</th>
<th>min_axis</th>
<th>ratio_x/y</th>
<th>ratio_z/x</th>
<th>ratio_y/z</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.14</td>
<td>0.98</td>
<td>1.08</td>
<td>0.52</td>
<td>0.33</td>
<td>0.28</td>
<td>0.40</td>
<td>0.14</td>
<td>0.01</td>
<td>0.87</td>
<td>1.07</td>
<td>1.33</td>
</tr>
</tbody>
</table>

(t(19)<-4.24, p<.01). This means that the new features contributed to improve the F-measure. Therefore, we analyzed the contribution of each feature based on information gain (IG). IG is commonly used in feature selection, where the gain of information provided by a particular feature is calculated by subtracting the conditional entropy with that feature from the entropy under random guess [44]. So, a more informative feature has a higher IG. Table 7 presents IG of each feature obtained from Sensor 7. In the table, ratio_y/z is the most informative among the 12 features. This is because the features raw_x and raw_y are also quite informative. The raw signals as well as the ratios between the two axes indicate the sensor tilt more directly than the other features. The z and y axes are effective to characterize rotation around the forearm, while the x axis contributes to represent bending and stretching of the elbow in the vertical direction. Rotation of the wrist affects the visibility of notification due to occlusion. So, we consider that such rotation-sensitive axes are more informative.

6. Conclusion

In this article, we proposed a method of identifying visible position on the wrist from ten candidate positions based on the forearm posture. A prototype system with ten full-colored LEDs and ten tri-axis accelerometers was implemented for demonstrating and evaluating the functionality of choosing the most visible LED on the wrist in different postures. A study with 20 users was conducted to see how well the selected LED matches with the users’ expectations. Also, further analysis was conducted for reducing the number of sensors while avoiding degradation of the selection performance is presented. The following results were obtained:

- The LED selector component could identify the most visible position among the ten candidates, aligned side-by-side at equal intervals and wrapped around the wrist, with an F-measure of 0.676 under user-independent evaluation of LOPO-CV using SMO. Thirty (10 sensors × 3 axes) raw acceleration signals were used as classification feature.
- The most visible position and an adequate position were selected at rates of 35.9 % and 94.9 %, respectively.
- Comparing the system’s judgments and the user expectations, 89.9 % of the system’s judgments fall within ±1 LED (± 18 mm) of what the users consider to be “the most visible position”, with the average gap between them being +0.73 LED (+13.14 mm).
- Information from the sensors located on the inside of the wrist individually yield a performance equivalent to that with all the ten sensors (F-measure of 0.681).
- New features mean, variance, max/min/argmax/argmin and ratios were introduced, in addition to the raw acceleration signals. Among these, the ratios between the two axes achieved the largest information gain. Rotation-sensitive axes, namely y and z, were found to contribute towards the system’s decision.

In this research, we assumed that the information is presented after a stable time period of at least 200 ms (10 samples at 50 Hz), so that the system can capture the forearm posture through the components of gravitational force. Though this assumption applied to a wide range of activity in our daily lives, there are many other activities where the user’s hand moves more quickly, for example while wiping a table, talking with hand gestures, and so on. We consider that it might be possible to utilize the stable axis as a feature for regular motions such as walking and wiping the table. For example, the stable axis during wiping the table is the z (vertical direction) axis while the directions of x and y axes change frequently. So, the z axis can provide reliable information to estimate the visible area on the wrist. By contrast, for irregular motions accompanied by rotation, such as hand gestures during talking, a gyro sensor can be used to track the posture of the display device in a continuous manner.

The results presented here were obtained with discrete LEDs tied around the wrist. In the future, we plan...
to implement more advanced version with a flexible and continuous display surface [30, 33] that can display more information like a particular application's icon with status information. In such a case, our method can be applied to a situation where the continuous display has the same number (10) and the same size (18 mm) of representative display areas as the current LED version. Two or three neighboring areas can be concatenated to display a larger piece of information.

Acknowledgment

This work was supported by JSPS KAKENHI 15K-00265.

References


Forearm Pose-aware Visual Notification for Whole Wrist Circumference Device


