Ambient Sensing Chairs for Audience Emotion Recognition by Finding Synchrony of Body Sway

Ryo Wataya, Daisuke Iwai, Kosuke Sato
Graduate School of Engineering Science, Osaka University
Machikaneyama1-3, Toyonaka, Osaka, 560-8531, Japan
Email: {wataya, iwai, sato} @sens.sys.es.osaka-u.ac.jp

***Abstract*—The authors aimed to develop a practical system to recognize group emotion without constraining users. This paper introduces a method to recognize the emotion of an audience attending a presentation by the ambient sensing of body sways or synchrony of multiple users. We developed ambient sensing chair, which employs wireless sensors to determine the center of force of seated users and a methodology for recognizing individual and group-acting-out behaviors. We designed an experiment to examine the relationship between audience body sways and emotions in a natural situation. Subjects watched five technical presentations, each of which consisted of 20 slides that switched automatically every 20 [s] and responded to questionnaires about their subjective impressions of each slide in each presentation. As a result of comparing subjects' acting-out behaviors and their synchrony to subjective impressions, synchrony could be well correlated to boredom. Our system was able to identify boredom at a recognition rate of more than 80%.**

I. INTRODUCTION

Emotion recognition research aims to provide core technologies for use in interactive systems. The majority of this research focuses on the recognition of a user's emotional state by using cameras, microphones, and biological sensors.

The use of cameras or microphones to recognize facial expressions and vocals tones is a typical approach[1]. This methodology is limited as these sensors are unable to function effectively in dark places or noisy environments. Therefore, the practical application of research employing this methodology is limited to theoretical experimentation. The addition of biological sensors, most of which are contact sensors[2], allows for broader practical applications. However, contact sensors introduce their own unique limitations: the constraint of a user's body and consciousness of the sensing system. As humans often form groups in daily life, there would also be a need for group emotion recognition, which few researchers have focused on. It is necessary to consider the difference between the simple average of each individual user's emotions and a group emotion. A specialized methodology is required to recognize group emotion.

The authors aim to develop a practical system for the recognition of group emotions that can be applied in various environments without introducing user constraint. This research consists of the following three approaches: sensing body sway and recognizing emotion, ambient sensing with a chair, and synchrony and group emotion recognition.

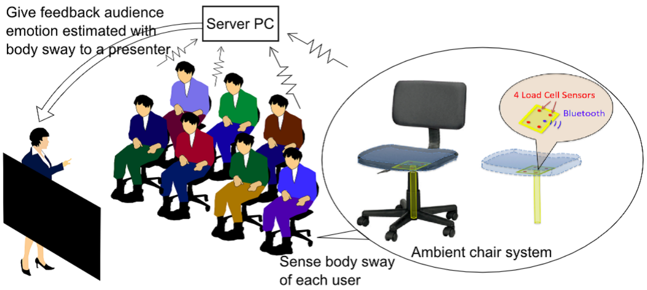

Fig. 1. System configuration: ambient sensing chairs equipped with four load cells sense body sways of multiple users without constraint and transfer the sensor data wirelessly to a PC. The system recognizes audience emotions and provide feedback to the presenter.

A. Sensing Body Sway and Emotion Recognition

Although humans often recognize the emotional states of others by vocal cues, body sway also provides important recognition cues. Body sway is observable through the use of force sensors. Therefore, monitoring the body sway could contribute to emotion recognition in environments where cameras and microphones are not appropriate.

B. Ambient Sensing with Chair

Most body sway research involves the use of chairs equipped with pressure or force sensors. Such systems are problematic because they are dependent on a high number of sensors. An early posture recognition study[3] used a chair equipped with a system of 4032 force sensors on the seat and back surfaces of the chair. Generally, such systems were not appropriate for widespread practical application because the implementation cost was too high and the high number of sensors resulted in users being conscious of the system in contradiction to our purpose. Recent studies have addressed these issues. Through the optimization of sensor placement, the number of required sensors was reduced to 19[4]. To improve upon this, the authors developed a chair system that uses only four pressure sensors, discretely positioned at the bottom of the seat of the chair. These chairs sense the center of pressure of a seated user without constraint and wirelessly transmit data to a server via Bluetooth.

C. Synchrony and Group Emotion Recognition

It is difficult to estimate emotion using only body sway data because body sway includes physiological reactions and expressions of emotion. In this study, the authors have focused on group reactions; the same reaction occurring simultaneously among multiple users. It is hypothesized that group reaction, which is a part of synchrony, would indicate shared emotional responses. Synchrony is defined as the dynamic and reciprocal adaptation of the temporal structure of behaviors between interactive partners[5]. For example, one user's body sway corresponding to another user's body sway would be a synchrony. Synchrony does not indicate physiological reactions but emotional reactions. This study aims to recognize audience emotion by sensing synchronous body sways.

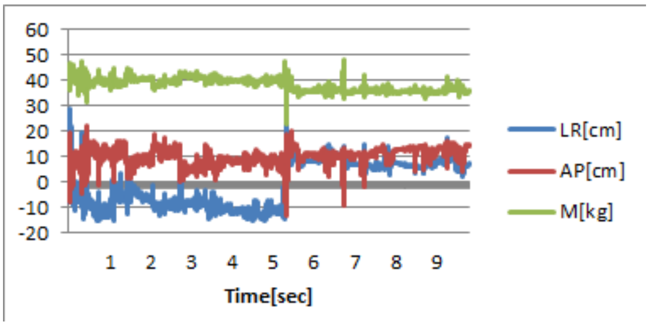

Fig. 2. Raw Data: Sensor data is translated into three dimensional data: LR[cm] (the center of force in the left-right direction), AP[cm] (the center of force in the anteroposterior direction), and M[kg] (load to the system).

II. METHODOLOGY

Ambient sensing chair system was developed to sense body sway without constraining a seated user[6]. The ambient sensing chair communicates with a server via Bluetooth, as shown in Fig. 1. This section explains the structure of the ambient sensing chair and the recognition methods for sensing body sway and the corresponding synchrony of users.

A. System Structure

The ambient sensing chair was constructed from an ordinary office chair. It has four load cells that are used for determining the force caused by a seated user's movements. It is sufficiently sensitive to detect subtle motions, such as a user lightly touching his/her nose.

The sensor data is transferred wirelessly to a PC via Bluetooth at a sampling rate of 40[Hz]. Six ambient sensing chairs are connected to the PC simultaneously. The sensor data, as shown in Fig. 2, is referred to as raw data in this article.

Each load cell has different output characteristics. Each ambient sensing chair is calibrated to sense the center of pressure and pressure load. Output data also includes noise caused from a sensor error; a median filter is used to remove noise, whose filter length is five sampling (0.125[s]).

B. Recognition Method

Acting-out behavior, which refers to a user's evoked behavior, and synchrony are detected by ambient sensing chairs. The method for detecting the acting-out behavior and the threshold and the recognition method for synchrony are explained below.

Initially, three-dimensional raw data is extracted for a particular width time Δ[s] window. The standard deviation (SD) of data within the width time window is calculated, as shown in Fig. 3. A high SD indicates that a seated user's center of pressure changes considerably or that the acting-out behavior is evoked. Here the acting-out behavior is defined as a SD over a certain threshold and will be labeled as *Acting-Out*. The duration of the acting-out behavior is plotted, as shown in Fig. 4. If no acting-out behavior is detected, the behavior is labeled as *Still*.

Determining the SD threshold is problematic. A seated user's body sway data includes clearly visible body sway, ambiguous or invisible body sways, and fluctuations caused by physiological reactions. It is difficult to divide the SD data into these distinct elements. Therefore, in this article, the threshold level is set as the minimizing error rate between Acting-Out and Still and a ground truth value manually assigned by an observer who watched a video of the experiment described in Section III. This process is applied to multiple seated users to recognize synchrony.

When a group of people watch television or a presentation, a phenomenon is observed in which people respond simultaneously or by a chain reaction. In this article, the phenomenon is defined as synchrony; if greater than or equal to n out of N seated users register labeled data within 2δ[s] of a particular time, then the system determines that synchrony has occurred.

Moreover, for each time period detected as synchrony, some users are labeled as *Involved-Sync.* if they are involved in evoking the synchrony. Those who are not involved are labeled *Not-Involved-Sync*. That is, in each time period detected as evoked synchrony, all users are labeled as either *Involved-Sync.* or *Not-Involved-Sync*. Synchrony is not necessarily evoked by all users. However, considering all users as a group, the internal states of some users who are not involved in a synchrony or who do not evoke the acting-out behavior would be changed because they are not totally independent of users who are involved in synchrony. For each time period and for each user, one label (*Acting-Out*, *Still*, *Involved-Sync*, *Not-Involved-Sync*) are assigned.

In the experiment described below, Δ was set as 1, δ as 1, N as 6, and n as 3.

III. EXPERIMENT

We conducted an external laboratory experiment to gain a deeper understanding of how body sway, resulting from a changed internal state under natural conditions, was evoked in users watching and listening to technical presentations in which slides are used. An additional purpose of the experiment was to examine the recognition rate of audience emotion by sensing their body sway, i.e., to determine if there is a relationship between a user's body sway and his/her emotions.

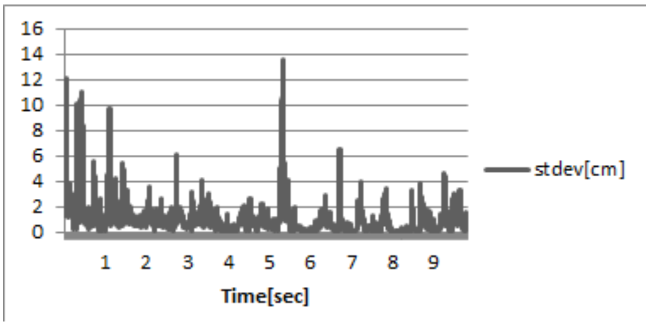

Fig. 3. Standard Deviation (SD) Data: High level of SD means that the center of pressure of a sitting person changes widely or that acting-out behavior is evoked.

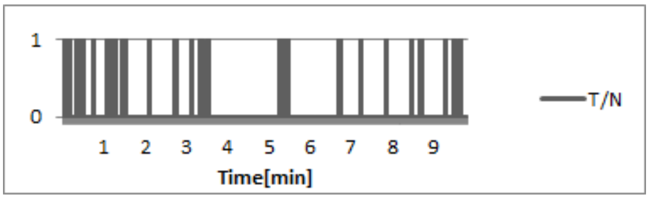

Fig. 4. Acting-out Label: The plotted periods are labeled as *Acting-Out*. (if no acting-out behavior is detected, the behavior is labeled as *Still*).

If users are aware that they are participating in an experiment designed to sense body sway, their reactions could be different from those in natural conditions. Since the goal was to recognize genuine emotion, it was important to design the experiment such that the subjects were not aware of the particular experimental conditions. In a previous study[7], the audience response tool (ART), the portable audience response facility (pARF), and the moment-to-moment index of eye movement were introduced as methods to evaluate audience emotion. However, the equipment involved with these methods would influence a subject's body sway and internal state. In contrast, our experiment has unobtrusive equipment. Subjects listened to a presentation while seated on ambient sensing chairs. The presentation was accompanied by slides that changed automatically at predetermined intervals, which allowed for the matching of sensor data with a particular slide. After the presentation, the participants completed a questionnaire designed to capture their subjective impressions of each slide. We analyzed the responses by comparing subjective impressions to the corresponding labels for body sway and synchrony and examined the relationships.

A. Experimental Method

1) Hypotheses: When listening to a presentation, agreeability, interest, and boredom are particularly important emotions. The authors hypothesized that affirmative nodding, not moving, and fidgeting were types of body sway that could correlate to agreeability, interest, and boredom, respectively. The ambient sensing chair system detected these three types of body sway, which were compared to three corresponding emotional responses ascertained from the questionnaires:

- **Agreeability Level**: How agreeable did a subject feel?
- **Interest Level**: How interested did a subject feel?
- **Fun Level**: How fun level did a subject experience?

2) Design: The experimental design needed to allow the researchers to analyze the relationships between agreeability, interest, and boredom, and acting-out behavior, or synchrony. The design needed to contend with two problems. Subjects are not always definitively aware of their emotional state and have difficulty accurately recalling how their emotions corresponded to the slides. As well, even if there is a strong relationship between body sway and emotions, there would be a time lag between the onset of a particular emotion and an acting-out behavior. To solve these problems, a small change of stimuli in a certain time was considered as a set of one. We also compared the emotional responses indicated in the questionnaires to the sensor readings to detect the acting-out behavior or synchrony for each particular time period.

Six subjects were seated in ambient sensing chairs that were lined up in a straight row. The subjects listened to a technical presentation that was accompanied by 20 slides that changed automatically every 20[s]. The ambient sensing chairs detected whether each subject displayed acting-out behaviors in each 20[s] interval. Soon after the presentation, subjects responded to a questionnaire in which they were asked to assign three subjective impressions (agreeability, interest, and fun) to each slide. This process was repeated five times, and the subjects listened to five different presentations.

We compared the subjective impressions to the detected acting-out behaviors and corresponding synchronies occurring within the interval for each subject. For instance, we compared the distribution of subjective values in the condition of the acting-out behavior to the condition of still.

3) Presenters: Five presenters were commissioned and each prepared a presentation. The presenters were native Japanese-speaking university students between 18 and 19 years. The presenters prepared and rehearsed a technical presentation about their student activities. All presentations were 6 minutes 40 seconds long.

4) Questionnaire: The questionnaires presented slide images with associated keywords and asked the subjects to assign a rating that corresponded to their emotional impression (agreeability, interest, and fun) of each slide. We employed a five-point Likert scale from -2 to +2.

5) Subjects: Six subjects (3 males and 3 females) were randomly chosen from the participants who visited this presentation conference. They were native Japanese-speaking university students between 18 and 22 years.

6) Procedure: Initially, the six subjects were seated on the ambient sensing chairs, which were lined up at 60[cm] intervals in the traverse direction and were given a brief description of the purpose and procedure of the experiment. After the introduction, they practiced responding to the questionnaires. The experiment was conducted in a conference center. Subjects listened to five presentations and responded to five corresponding questionnaires in a natural condition. Other participants also responded to the same questionnaires

Fig. 5. Experimental Setup: Six subjects sitting in ambient sensing chairs in the front row.

in the same location. The questionnaires were administered soon after the technical presentations finished.

After the experiment, the subjects were asked whether they had discerned that the ambient sensing chairs were not ordinary chairs. None of the subjects had detected that the ambient sensing chairs had been monitoring their body sway. They were informed that their physiological responses had been monitored and recorded by the ambient sensing chairs. All subjects gave permission to use and publish the gathered data.

Figure 5 is a photograph of the experimental environment. The six subjects were seated in ambient sensing chairs in the front row. The other attendees were seated in normal chairs. The ambient sensing chairs were well integrated and monitored the subjects wirelessly. The subjects were not constrained in any way by the monitoring devices.

B. Data Analysis

As explained previously, each time period was labeled as *Acting-Out* or *Still*. In addition, some periods detected as Acting-Out were labeled as *Involved-Sync* or *Not-Involved-Sync*. Therefore, each period is classified as one of *Acting-Out*, *Still*, *Involved-Sync,* or *Not-Involved-Sync*. The evaluation value $E_{s,L}$ of a period labeled as L for subject s is defined below:

$$E_{s,L} = Q_{s,L} - A_s,$$

where $Q_{s,L}$ is defined as the average of the Likert scale ratings by subject s in the period L, and A_s is defined as the average of the Likert scale ratings by subject s in all periods. If $E_{s,L}$ is a positive number, then subject s felt more positive emotion than normal in period L. We used the evaluation value $E_{s,L}$ to analyze the experimental data.

C. Results and Discussion

Table I shows the occurrence of synchrony. It can be seen that, on average, subjects evoked synchrony 2.2 times during the 20 slide presentation and that synchrony was probably evoked in the first (1st(F)) and last (20th(L)) slide.

Figure 6 shows the relationship between the evaluation values, $E_{s,L}$ and the detected labels for each subject's acting-out behavior unrelated to synchrony. Each bar indicates the range of the evaluation values, $E_{s,L}$, for subject s in period L for each emotion. The top row contains the maximum evaluation values, $\max\{E_{s,L}\}$, and the bottom row contains the minimum evaluation values, $\min\{E_{s,L}\}$; $\times$ indicates the values of average $\bar{E_L}$.

In Fig. 6, all bars extend above and below 0 (i.e., the top of the bar is positive and the bottom is negative). This result means that when the labels were *Acting-Out* or *Still*, some subjects felt positive emotions compared to their normal state, and other subjects felt negative emotions. In other words, from the viewpoint of the existence or non-existence of a single user's acting-out behavior unrelated to synchrony, a common emotional response was not observed.

Figure 7 shows the relationship between the evaluation values, $E_{s,L}$, and the detected labels for each subject's synchrony. This graphic representation is similar to that show in Fig. 6. The first four bars from the left, agree and interest, also extend above and below 0. This result means that the common response of agreeable emotion and interested emotion was not observed to be related to involvement or non-involvement with synchrony.

On the other hand, the two bars on the right, fun and bored, are both below 0. In other words, the tops of the bars, i.e., the maximums of evaluation value, E, are negative. This means that even if both *Involved-Sync.* (In-Sync.) and *Not-Involved-Sync* (N-In-Sync.) labels were examined, no subjects indicated a positive emotional response compared to their normal state during the synchrony time periods. From the result, synchrony could be well correlated with boredom, and boredom could be estimated by synchrony.

To confirm the feasibility of the above assumption, the recognition rate of boredom synchrony was calculated. The recognition rate is defined as the ratio of the negative value of the number of periods evaluated to the number of periods labeled as *Involved-Sync.* or *Not-Involved-Sync*. As the rate was greater than 80%, our system was able to recognize audience boredom with high accuracy.

On the other hand, the two bars from the right, that is fun (or bored) level, were under the horizontal line. In other words, the tops of bars, that is the maximum of evaluation value E were negative. This means that even if both *Involved-Sync.* (In-Sync.) and *Not-Involved-Sync.* (N-In-Sync.) labels were focused, all subjects felt negative emotion of fun compared to their normal in the time periods labeled as synchrony. From the result, synchrony could be well correlated with bored emotion, and bored emotion could be estimated by detected synchrony.

In order to confirm the feasibility of the above assumption, the recognition rate about bored emotion was judged by the synchrony was calculated. In this article, the recognition rate is defined the ratio of the number of periods evaluated as negative value to the number of periods labeled as *Involved-Sync.* or *Not-Involved-Sync*. As the rate was more than 80 %, our system is able to recognize audience bored emotion with a high degree of accuracy.

TABLE I
SYNCHRONY OCCURRENCE

Presen.	Slide	Share of Sync.	s.1	2	3	4	5	6
1st	1st(F)	6	o	o	o	o	o	o
1st	9th	4	o	o	o		o	
1st	20th	3	o			o	o	
2nd	20th(L)	5	o	o		o	o	o
3rd	1st(F)	6	o	o	o	o	o	o
3rd	18th	3	o		o			o
3rd	20th(L)	6	o	o	o	o	o	o
4th	1st(F)	5	o	o	o	o		o
5th	1st(F)	3	o	o		o		
5th	13th	4	o			o	o	o
5th	20th(L)	3	o	o				o
		Total	11	8	6	8	7	8

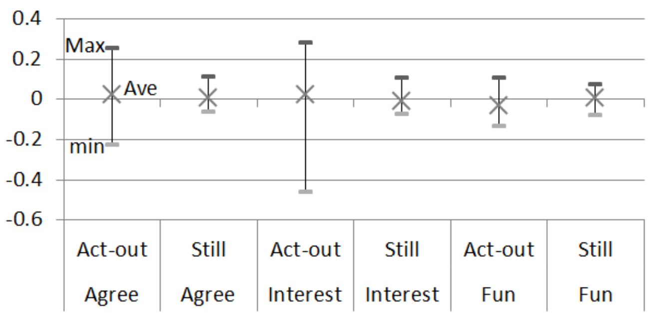

Fig. 6. Relationship between the evaluation values E and detected labels for each subject's acting-out behaviors unrelated synchrony. The two bars from the right, that is fun (or bored) level, were under the horizontal line.

IV. PROSPECTIVE APPLICATION

These results indicate the possibility of developing systems that give effective feedback about audience boredom levels to a presenter in either online or offline environments. For example, a prospective real-time presenter support tool that advises the presenter to change the topic, or their way of speaking. A refining tool that enables the presenter to improve their presentation by comparing the audience's post-presentation comprehension with the data collected by the ambient sensing chair system could also be developed. Additionally, this knowledge allows for the possibility of developing emotion feedback tools, such as remote teaching-support systems and emotion-based feedback systems for producers of television shows and commercials, movies, and live theater. Such feedback systems could allow producers to assess their productions beyond the limitations of traditional feedback, such as written questionnaires.

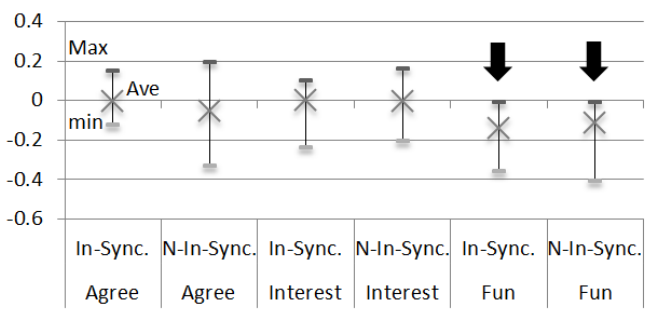

Fig. 7. Relationship between the evaluation values E and detected labels for each subject's synchrony. This means that even if both *Involved-Sync.* (In-Sync.) and *Not-Involved-Sync.* (N-In-Sync.) labels were examined, all subjects felt negative emotion of fun compared to their normal state in the time periods labeled as synchrony.

V. CONCLUSION

First, an ambient sensing chair system was developed using an ordinary office chair and devices to sense the center of force of a seated user. Second, experiments involving live presentations were conducted to investigate the relationship between audience members body sways and their emotions. The results of comparing the evaluated body sway and corresponding synchrony among audience members showed that the system could reliably identify synchronous boredom. The recognition rate of boredom was greater than 80%, indicating that the identification of emotional synchrony may be a valuable feature quantity. Future work will confirm the effectiveness of the ambient sensing chair system and synchrony in other situations, such as watching movies and TV.

REFERENCES

[1] Glenn I. Roisman Zhihong Zeng, Maja Pantic and Thomas S. Huang.: ”A survey of affect recognition methods: Audio, visual, and spontaneous expressions,” IEEE Transaction on Pattern Analysis and machine Intelligence, 31, pp.39-58, 2009.

[2] Kim J. and Andre.: ”Emotion recognition based on physiological changes in music listening,” IEEE Transactions on Pattern Analysis Machine Intelligence, 30(12), pp.2067-2083, 2008

[3] Hong Z. Tan, Lynne A. Slivovsky, Alex Pentland: ”A Sensing Chair Using Pressure Distribution Sensors,” Transactions on Mechatronics 6(3), pp.261-268, 2001

[4] Bilge Mutlu, Andreas Krause, Jodi Forlizzi, Carlos Guestrin, Jessica Hodgins: ”Ro- bust, Low-Cost, Non-Intrusive Sensing and Recognition of Seated Postures,” Human- Computer Interaction Institute. Paper 42, http://repository.cmu.edu/hcii/42, 2007

[5] Delaherche, E. and Chetouani, M. and Mahdhaoui, A. and Saint-Georges, C. and Viaux, S. and Cohen, D.: ”Interpersonal Synchrony : A Survey Of Evaluation Methods Across Disciplines,” IEEE Transactions on Affective Computing, pp.1-20, 2012

[6] Shohei Matsumoto, Daisuke Iwai, Kosuke Sato: ”Sedentary Body Sway Sensing System for User Behavior Recognition,” SDPS Kansai sector conference, C-27, 2011 (in Japanese)

[7] Catherine Stevens, Renee Glass, Emery Schubert, Johnson Chen, Heather Winskel: ”METHODS FOR MEASURING AUDIENCE REACTIONS,” The inaugural International Conference on Music Communication Science, pp.155-158, 2007