Community Detection in Real Communication Using Body Sway and Voice Information with Ambient Sensing Chair

Ryo Wataya, Daisuke Iwai and Kosuke Sato

***Abstract*— The authors aim to develop a community detection system without constraining users. This paper introduces a method to detect communities in which users are engaging in a conversation with each other, by an ambient sensing of body sways and their synchrony between users. We apply an improved *ambient sensing chair*, which had been developed in our previous study. The system employs wireless force sensors to measure the center of pressure. By using measured speaker/listener information and synchrony between any users' body sways, it is confirmed to detect the community information in experiment.**

I. INTRODUCTION

Human can detect a target community constructed from multiple users, and take appropriate action adapted to the community. Considered an example about a clerk in a restaurant. A group customers who visit the restaurant together will expect to be taken orders and to be served each dishes in the same time. If the clerk did not understand community information about the customers, he/she could not serve with higher customer satisfaction. Basically, end user services need to know an information on customer communities for realizing high quality user service.

On the other hand, interactive systems have recently been developed by many researchers. Some systems can change their operations adapted to a specific user, only in the case where the number of target users is one, by learning the user's property such as his/her habit and behavior. However, in the case of community constructed from multiple users, almost of these systems operate uniformly to each community.

To develop interactive systems or services for communities in the real world, the community detection is one of key issues to be solved; that is a problem about finding a speaker and listeners who involve a conversation.

Human can detect most of information on communities from the visual information such as position, eye direction, audio information such as back-channel feedback from a listener to a speaker, and so on. It is reasonable configuration with consisting of many cameras to realize automatic community detection. In the case of camera, it may invade user's privacy and has many occlusions in its sight.

In this paper, a situation where group users are sitting is assumed. In this condition, the authors focus a sensing method without introducing user constraint of his/her movement and consciousness. This sensing method is called as ambient sensing[1]. The other sensing methods, which constrain users, can make users' behavior change and invade their natural actions. Therefore, the ambient sensing will be applied for realizing Ambient Intelligence.

Graduate School of Engineering Science, Osaka University
Machikaneyama-1-3, Toyonaka, Osaka, 560-8531, Japan
{wataya,iwai,sato}@sens.sys.es.osaka-u.ac.jp

People engaging a conversation on chairs are likely to convey verbal and nonverbal behavior (e.g. nod) between a listener and the speaker in their communication. Basically, the correlation may observed sharply between speaker's verbal information and listener's movement from our observation[9].

In this paper, feasibility of our system for community detection is examined in two cases: dyadic communication and multi-party communication. This way is based on psychological study method, in which dyadic and multi-party communication are dealt with total differently. For realizing community detection, the users are sensed ambiently with special developed chairs concealed-fixed to microphone for obtaining their verbal information and load cells for obtaining their body sway information.

Although position information and visual information could achieve our goal easily, no any location sensors and cameras are used in this research, because this research is also aimed at for the case of talking with distant-existing users.

II. RELATED WORKS

A. Community Detection

Researches about community detection are classified roughly into detection from graph[2] and detection from social network[3]. Web data is used in almost all of the researches, but there are few research about group estimation using a real communication information.

The former has been researched for some decades. Traditional methods for detecting community are Graph Partitioning, Hierarchical Clustering, Partition Clustering, and Spectral Clustering. Recently, applied methods such as technique in the case of overlap with some communities are studied.

B. Psychological Findings in Communication

Some communication channels such as sound with affect expression, body motion, pose, mannerism, the number of utterance, utterance volume, basic vocal frequency, facial expression, and so on, are synchronized in a conversation[4]. The body sway synchrony has been applied to estimate affect recognition of sitting users by the authors in the previous paper[5]. The communication channels are divided into three groups: body motion information, vocal information, and

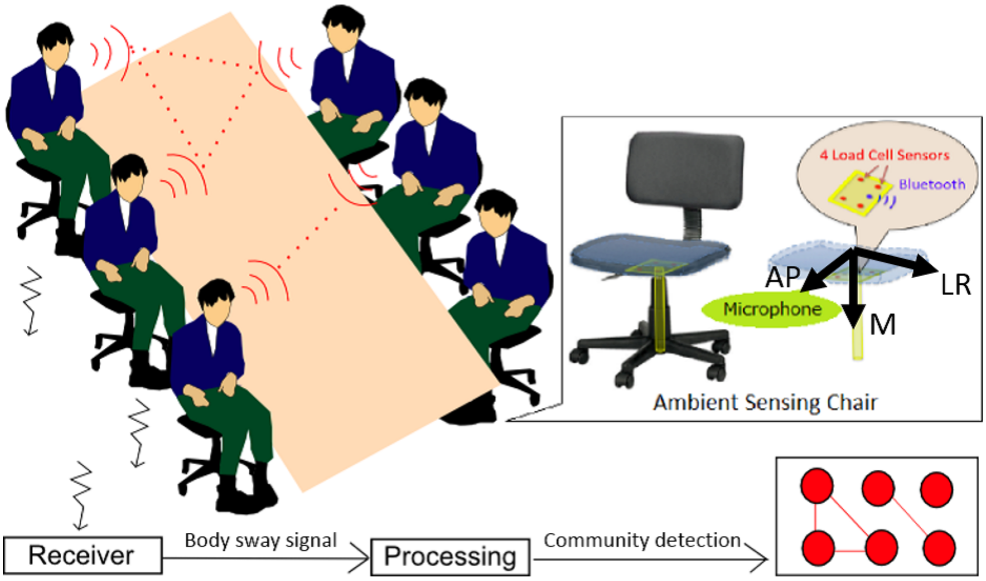


Fig. 1: System Configuration: Ambient sensing chairs equipped with four load cells and a microphone sense body sways and voices of multiple users without constraint and transfer the sensor data wirelessly to a server. The system detects communities in which users are engaging in a conversation with each other and visualize the detected result as a graph representation.

facial expression. Sensing of these information enables us to be given the cues about which users form a group[6].

C. Sensing Method of Sitting Person's Body Sway and Utterance

Most researches about sensing body sway or posture involves the use of chairs equipped with force sensors. Such systems are problematic because they are dependent on a large number of sensors. An early posture recognition study used a chair equipped with 4032 force sensors on the seat and back surfaces of the chair[7]. Generally, such systems were not appropriate for widespread practical application because the implementation cost was too high and the large number of sensors resulted in users being conscious of the systems 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[8].

To improve upon this, the authors developed a chair system in the previous papers[5], [9], [10], called *ambient sensing chair*, 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. In this paper, the authors add a microphone to the ambient sensing chair to enable user's voice sensing.

D. Our Target

The authors hypothesize by using multiple improved ambient sensing chairs, sensed body sway and voice data would enable community detection in the real world communication. In this paper, community detection system is implemented by using synchrony information about body sway of multiple users which are recognized ambient sensing chairs.

III. METHODOLOGY

A moderately priced chair system, which we refer to as the ambient sensing chair, was developed to sense body sway and voice without constraining a seated user. 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 between multiple 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 movement and a microphone for determining existence or non-existence of user's utterance. It is sufficiently sensitive to detect subtle motions, such as a user lightly touching his/her noise.

The sensor data is transferred wirelessly to a server via Bluetooth at a sampling rate of 40[Hz]. Six ambient sensing chairs are connected to the server simultaneously.

Each load cell has different output characteristics. Each ambient sensing chairs is calibrated to sense the center of pressure and pressure load. As an embed sensor outputs signals contaminated with spike noises, a five tap median filter (window width: 0.125[sec]) is applied to remove ones.

B. Recognition Flow

This system operates as the flow as shown in Fig. 2.

Input data of body sway is converted into three-dimensional data as shown in Fig. 1: AP[cm] is the center of force in the anteroposterior direction, LR[cm] is the center of force in left-right direction, and M[kg] is the load to the system. The three-dimensional data is cut by time window at

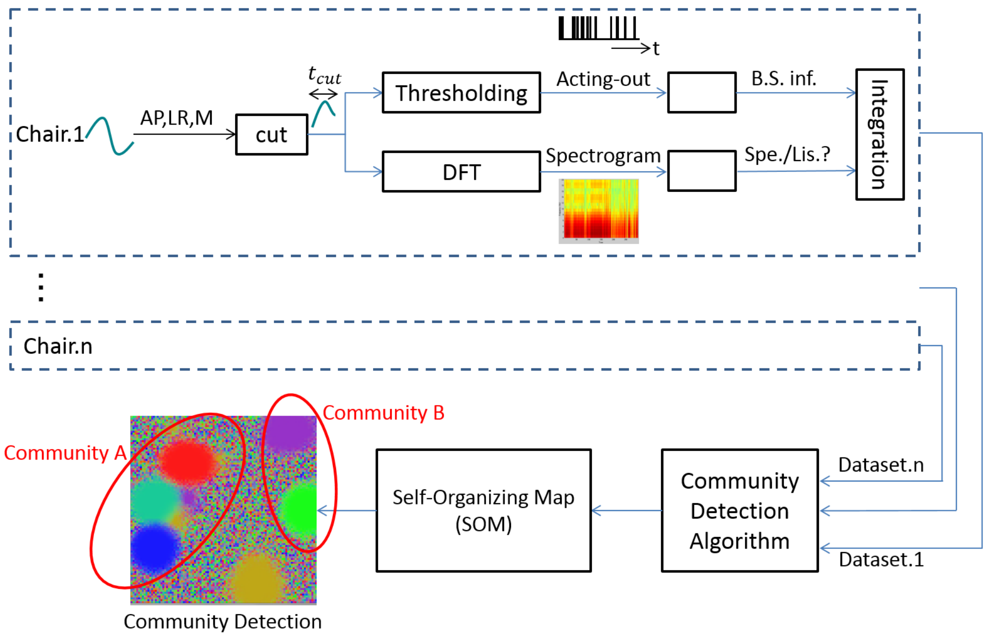


Fig. 2: Processing Flow: Input body sway data (AP, LR, M) of a subject (chair) is cut by time window at the size of t_{cut}. Cut data is set as threshold, which gives body sway information (B.S. inf.) about when the user did acting-out behavior. Simultaneously, calculated discrete Fourier transform (DFT) of cut data gives spectrogram, which could be a feature for detecting whether the user is speaker (Spe.) or listener (Lis.). These information of multiple users are integrated for detecting community. The result about community detection is visualized with self-organizing map (SOM).

the size of t_{cut}. Cut data is classified by means of threshold processing, that gives body sway information. The information about when the user did acting-out behavior, is plotted called as acting-out label. Simultaneously, calculated discrete Fourier transform (DFT) of cut data gives spectrogram. This spectrogram data could be a feature for detecting whether the user is speaker or listener.

These information of multiple users are integrated for detecting community. The result about community detection is visualized with self-organizing map (SOM).

C. Speaker/Listener Detection

Body sway row data is processed with short-time analysis with hamming window (window size: T [sec], slide range: 0.125 [sec]). From the obtained DFT data, the spectrogram is calculated for obtaining time variation of amplitude at each frequency.

With time variation in frequency f[Hz], a particular user is estimated as speaker or listener in each time interval with threshold th[cm]. In this experiment, recognition rates R with three parameters (T, f, th) are examined and optimal these parameters will be examined.

D. Dyadic Community Detection

Correlation score between subject a (Sub.a) and subject b $E_{a,b}$ is defined as the below correlation function:

$$E_{a,b} = \Sigma_t W_c \Delta_t, \tag{1}$$

where W_c, Δ_t are defined according to three cases in Fig. 3. Case1 is strong correlation between Sub.a and Sub.b, case2 is weak correlation, and case3 is otherwise (no correlation). Each rectangle indicates acting-out label, which means time period when body sway arises. R is constant value and r is minimum time period between acting-out label of Sub.a and one of Sub.b.

E. Multi-party Community Detection

Every time when synchrony among any two and more subjects is recognized, links between corresponding nodes have been connected for t_{link} [sec] for visualization. Synchrony is recognized if two and more acting-out behaviors of subjects detected every Δ_t [sec] are lapped within t_{int} [sec]. The acting-out behavior is decided to be detected at the occurring frequency of p' [%], given from p-tile method. In this experiment, these parameters were set t_{link} as 180, Δ_t as 1, t_{int} as 1, and p' as 2.

These parameters were set based on the below: in this experimental set, the probability P of detecting synchrony in a minute between any two subjects chosen randomly is about 32[%], and probabilities P' (less than 3[%]) of detecting synchrony in a minute among any three and more subjects chosen randomly is much smaller than P. P and P' are dependent on Δ_t, t_{int}, and p' and these values were set for easier analysis and less noise. Thus, t_{link} was set appropriately for reducing error detection.

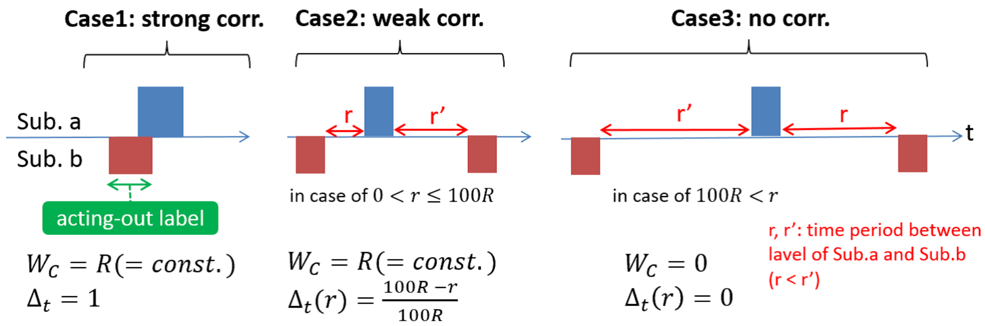


Fig. 3: Definition of Correlation Function: Parameters W_c, Δ_t in $E_{a,b} = \Sigma_t W_c \Delta_t$ are defined according to three cases in this figure. Case1 is strong correlation between Sub.a and Sub.b, case2 is weak correlation, and case3 is otherwise (no correlation). Each rectangle indicates acting-out label, which means time period when body sway arises.

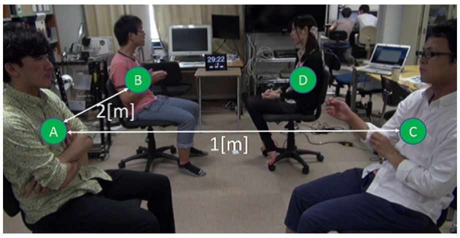


Fig. 4: Experimental Setup: Two subjects were seated on the ambient sensing chairs face-to-face at about 1[m] distance, and the other two subjects were seated in the same way, which were lined up at about 2[m] distance. A to D indicate subjects' symbols.

TABLE I: Role Turn in this Experiment: A to D indicate subjects' symbols. 1 to 6 indicate experimental order. Speaker1 and Listener1 talk with each other. Speaker2 and Listener2, similarly.

#	Speaker1	Listener1	Speaker2	Listener2
1	A	B	C	D
2	B	A	D	C
3	A	C	B	D
4	C	A	D	B
5	A	D	B	C
6	D	A	C	B

IV. EXPERIMENT1: COMMUNITY DETECTION FOR DYADIC COMMUNICATION

A. Experimental Method

1) Objective: Dyadic communication is defined as two users form a community and talk with each other in a same room in this paper. In a dyadic communication, exchange of communication, such as nod, behavior between a speaker and a corresponding listener is an only cues for community detection. It is effective to discriminate speakers from listeners before analyzing relationship of communication behavior, because the discriminated data decreases the number of candidate pairs for each subject.

Thus, the authors examine the feasibility of community detection using the cue of behavioral synchrony between a speaker and a listener, by comparison between using only behavioral synchrony information and using the information combined with detected speaker/listener information. In the comparison, correlation score originally defined in this paper is adopted.

2) Design: Two subjects were seated on the ambient sensing chairs face-to-face at about 1[m] distance, and the other two subjects were seated in the same way, which were lined up at about 2[m] distance, as shown in Fig. 4. A subject was given a speaker role and the other was given a listener role. The two subjects talked with each other about any topic for three minutes. The speaker just speaks and the listener listens to the speaker's utterance with back-channel feedback. Combinations of pairs were changed as shown in Table 1 and six combinations data were obtained.

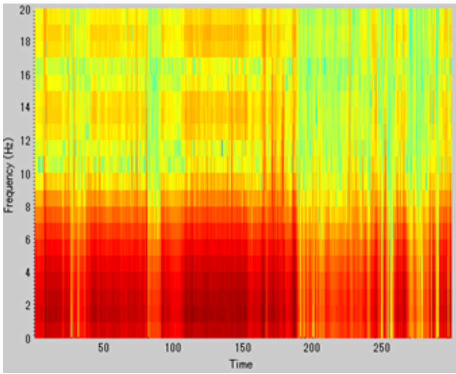
3) Subjects: 4 subjects (3 males and 1 females) were randomly chosen from our laboratory. They were native Japanese-speaking university students between 21 and 23 years.

In the only analysis of speaker/listener detection, other 12 subjects (11 males and 1 females) between 20 and 24 years are also chosen, for reduction of individual difference.

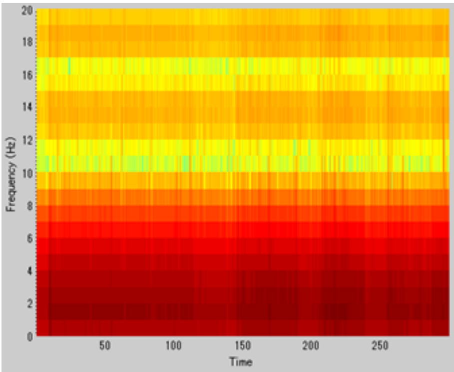
B. Results and Discussion

1) Speaker/Listener Detection: Figure 5 shows the results of spectrogram of body sway about a speaker and a listener. Horizontal line indicates time, vertical line indicates frequency, red indicates strong change, and green indicates weak change. From these figures, body sway of speaker is not steady contrary to one of listener. This means a listener is apt to hear a speaker voice with immobility and speaker is apt to utter with some gestures such as adaptors, regulators, illustrators, affect displays or emblems, classified by Ekman in 1969. Therefore, variance of a frequency in body sway could be feature quantity to recognize speaker or listener.

Next, the maximum recognition rate and the optimum parameters (T, f, th), which are defined in Chapter 2-C, were examined. When T was set as 5[sec], f as 1[Hz], and th as 5[cm], the recognition rate was maximum and the rate was 74 [%]. Most of incorrect periods stemmed from listener's laughing in the conversation.



(a) Speaker



(b) Listener

Fig. 5: Spectrogram of a Speaker(a) and a Listener(b): Horizontal line indicates time, vertical line indicates frequency,red indicates strong difference, and green indicates weak difference. Body sway of speaker is not steady contrary to one of listener.

2) Dyadic Community Detection: In Table 2, correlation score $E_{a,b}$ of part of experiments (only Experiment 1, 2, 3 and 4) defined in Table 1 was arranged. Take attention that definition of the correlation score was not symmetric property. The true combination about conversation pairs was shown in Table 1. For example, the correlation value of subject A in experiment1 to subject B (6.9) is higher than to subject C (2.5) or subject D (4.0). These values meant that the most strongly correlated subject to subject A was estimated to subject B, and the true combination is subject A and B in Table 2(a). Therefore, this result was correct.

From these results, the number of correct combination was 14 in total 24 (58 [%]). Combining this data with estimated speaker/listener data, the correct rate rose to 66 [%], because the data decreases the number of candidate pairs for each subject. This correct value was not still feasible in terms of feasibility in some practical applications, so voice information should be combined to improve the correct rate.

TABLE II: Experiment Results: Correlation score $E_{a,b}$, defined in Table 1, is arranged. For example, the correlation value of subject A in experiment1(a) to subject B (6.9) is higher than to subject C (2.5) or subject D (4.0). These values meant that the most strongly correlated subject to subject A was estimated to subject B.

(a) Experiment1

	A	B	C	D
A	/	6.9	2.5	4.0
B	4.5	/	1.6	9.5
C	3.0	1.6	/	1.6
D	6.9	11.4	3.2	/

(b) Experiment2

	A	B	C	D
A	/	3.6	5.4	2.7
B	3.4	/	2.7	1.3
C	4.8	2.1	/	3.7
D	3.3	1.3	3.4	/

(c) Experiment3

	A	B	C	D
A	/	4.8	6.4	7.3
B	4.3	/	2.6	2.3
C	7.3	5.9	/	5.0
D	7.1	2.5	3.9	/

(d) Experiment4

	A	B	C	D
A	/	1.3	0.0	4.7
B	1.3	/	2.1	3.6
C	0.0	0.0	/	3.2
D	3.2	2.0	0.0	/

V. EXPERIMENT2: COMMUNITY DETECTION FOR MULTI-PARTY COMMUNICATION

A. Experimental Method

1) Objective: Multi-party communication is defined as three and more users form a community and talk with each other in a same room in this paper. In multi-party communication, synchronization among users, especially listeners, is focused in some studies such as by Evangelist et al.[13]. Thus, the authors examine the feasibility of community detection using the cue of synchrony among listeners, by comparison between synchrony among listeners who are engaging in a same conversation and who are engaging in each different conversation.

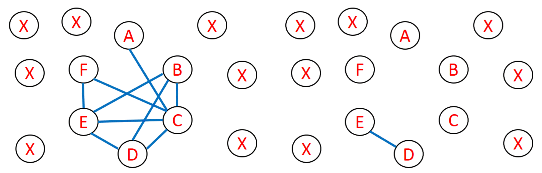
2) Hypothesis: The authors hypothesized the below:

- Synchrony between any two listeners is stronger than that between a speaker and any listener.
- Synchrony among listeners occurs even if their distance is a certain far.

Thus, synchrony among listeners, which is occurred without depending on their distance, is utilized for community detection in this experiment.

3) Design: A subjects was seated on the ambient sensing chair at about 80[cm] distance from a display. The subject watched a video in which a speaker talk to any listeners in the form of one way communication. The duration time of the video was about 7[min]. The subject was indicated to watch two types of videos as involving in a conversation with the speaker. This process was repeated six times in a same room, by changing subject.

4) Subjects: 6 subjects (5 males and 1 females) were randomly chosen from our laboratory. They were native Japanese-speaking university students between 21 and 23 years.



(a) True Combination (b) Wrong Combination

Fig. 6: Detected Community in Two Cases: (a)True case: Subject A to F in the figure indicate listeners who are engaging in a same conversation. All subjects were linked with each other. (b)Wrong case: Subject A to F are engaging in different conversation. Only subject D and E were linked by mistake.

B. *Results and Discussion*

Figure 6 shows as the part of results of community detection in this experiment. Figure 6(a) indicates detected community in the case of true combination, that is subject A to F in the figure are listeners who are engaging in a same conversation. Links started to be connected about 2[min] after the experiment began, and correct recognition had finished since about 5[min] were passed.

On the other hand, Figure 6(b) indicates detected community in the case of wrong combination, that is subject A to F in the figure are listeners who are engaging in each different conversation. Wrong link as shown in this figure was connected only one time between a couple of listeners. Thus, the recognition method for community detection was well conducted.

The number of synchrony among true combination was 5 to 9 times more than one among wrong combination. It means that the synchrony among listeners is an efficient feature for detecting users engaging in the same community.

VI. DISCUSSION

From the two experiments, we found that community detection for multi-party communication worked with higher feasibility than that for dyadic communication. The main reason is that synchronous feature quantity of multi-party communication was much stronger than one of dyadic communication in terms of the number of synchrony.

The ratio of synchrony occurred among users in the same community is much higher than one of synchrony occurred among users in different communities, in the case of that there are small number of communities where a lot of users belong to each community. Therefore, our systems would provide good estimation in the case. In the next step, the authors will examine the detail relationship between these numbers and the feasibility of estimation.

The authors are going to improve to develop system which can visualize estimated output momentarily, as using self-organizing map (SOM), shown in Fig. 1. SOM is the visualizing method, with which distance characteristics on multidimensional vector space is kept on projecting onto ones on low dimensional vector space (mostly one to three dimension). This method enables data on multidimensional vector space difficult to understand for us to visualize understandably on a display device.

Audio data should be utilized for community detection. Correlations among audio data sensed from microphones attached with ambient sensing chairs enables to richer cue for community detection such as their orientations and distances.

As future work, our system will be applied in the real world such as cafeteria in a university. Experiment in a natural condition will inspire numerous suggestions for possibility of our system.

VII. CONCLUSION

Community detection for multi-party communication is conducted correctly with developed ambient sensing chair, which enables to sense a sitting user's body sway without constraining his/her body and consciousness. And it indicates higher feasibility than one for dyadic communication, for the main reason of much stronger synchrony in multi-party communication than one in dyadic communication.

REFERENCES

[1] Weber, Werner, Jan M. Rabaey, and Emile HL Aarts, eds.: "Ambient intelligence", Springer, (2005).

[2] Santo Fortunato: "Community detection in graphs", Physics Reports, Vol. 486, pp. 75-174 (2010).

[3] S. Papadopoulos, Y. Kompatsiaris, A. Vakali, P. Spyridonos: "Community Detection in Social Media", In Data Mining and Knowledge Discovery, Vol. 24(3), pp. 515-554 (2011).

[4] Chika Nagaoka: "Mutual influence of nonverbal behavior in interpersonal communication", Unpublished doctoral dissertation <http://ir.library.osaka-u.ac.jp/dspace/bitstream/11094/5842/1/jjisp06_101.pdf>, Osaka University, (2003).

[5] Ryo Wataya, Daisuke Iwai, Kosuke Sato: "Sensing of Excitement Based on the Synchrony of Sitting Body Sway", IEEJ Transactions on Electronics, Information and Systems, Vol. 132(1), pp. 1-7 (2013) (in Japanese).

[6] M. Javier Klaicha, Paul G. Kinas, Susan N. Pedraza, Mariano A. Coscarella, Enrique A. Crespo: "Estimating dyad association probability under imperfect and heterogeneous detection", Ecological Modelling Vol. 222, pp. 2642-2650 (2011).

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

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

[9] Ryo Wataya, Daisuke Iwai and Kosuke Sato: "Ambient Sensing Chairs for Audience Emotion Recognition by Finding Synchrony of Body Sway", In Proceedings of IEEE Global Conference on Consumer Electronics (GCCE), pp. 29-33 (2012).

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

[11] Yano, Kazuo, et al.: "Measurement of human behavior: creating a society for discovering opportunities", Hitachi Review 58.4 (2009).

[12] Ara, Koji, et al.: "Healthcare of an organization: using wearable sensors and feedback system for energizing workers." Proceedings of the 16th Asia and South Pacific Design Automation Conference. IEEE Press, (2011).

[13] Evangelist Michael, Francez Nissim, Katz Shmuel: "Multiparty interactions for interprocess communication and synchronization", Software Engineering, IEEE Transactions Vol. 15(11), pp. 1417-1426 (1989).