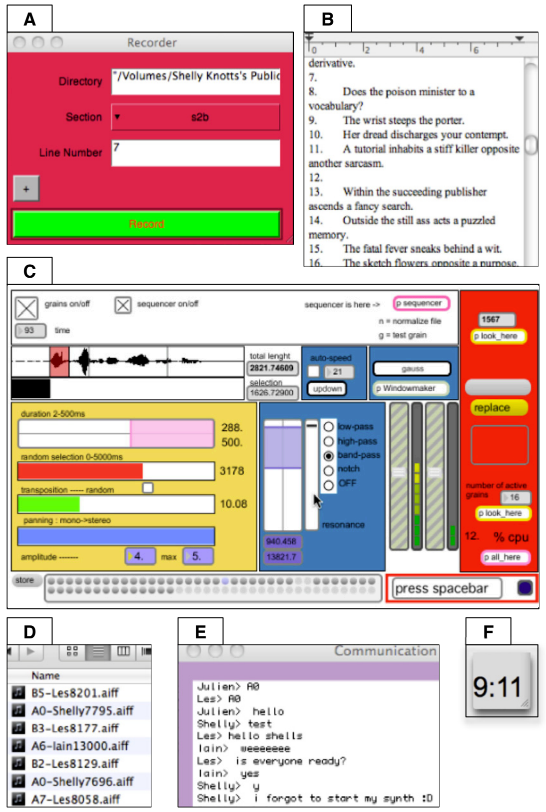

Figure 4. Elements of a Socially Constructed Performance System.

5. CONCLUSIONS

In the course of this paper, we have presented an ethnographic study of performance modalities in BiLE's Laptopera Act 2, identifying the stakeholders involved in each modality and considering how multiple modalities manifest themselves as socially constructed performance systems. Our investigation shows that whilst traditional roles of composer and performer are present within BiLE practice, the consensus driven nature of the group produces a dynamic set of orientations, which do not strictly define players' sole activities. Instead, interaction is perhaps best understood with reference to hybrid roles such as a) the *composer-performer*, reflecting the fact that the member who conceives of a piece is also involved in playing it, b) the *composer-designer*, reflecting the need for composers to design infrastructures which aid realisation of the piece by simplifying interaction, and c) the *performer-designer*, reflecting the fact that performers are required to design instruments which stand apart from each other. As these latter two roles show, design inheres within both composition and performance, but in service of different sociotechnical functions, representing the need for standardisation of particular modalities, whilst allowing for individuation in others. In addition, other modalities can be described as emergent, due to their reliance on tools or approaches that the group have developed as part of their wider practice.

Taken together, these findings reveal the essentially complex, socially constructed nature of musical interaction in BiLE practice. Here it is precisely the way in which the bounds of collaboration are negotiated anew - rather than their explicit formalisation - that acts as the primary driving force in the creation of new work.

6. REFERENCES

[1] "Bile manifesto." [Online]. Available: http://bilensemble.wordpress.com/manifesto/

[2] S. Beck, C. Branton, and S. Maddineni, "Tangible performance management of grid-based laptop orchestras," *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 207–210, 2011.

[3] G. Booth and S. Gurevich, "Collaborative composition and socially constructed instruments: Ensemble laptop performance through the lens of ethnography," in *12th International Conference on New Interfaces for Musical Expression, Ann Arbor, Michigan, USA*, 2012.

[4] G. Button, "The ethnographic tradition and design," *Design Studies*, vol. 21, no. 4, pp. 319–332, Jul. 2000.

[5] S. Cottrell, *Professional music-making in London: ethnography and experience*. Ashgate Publishing, Ltd., Sep. 2004.

[6] A. Crabtree, S. Benford, C. Greenhalgh, P. Tennent, M. Chalmers, and B. Brown, "Supporting ethnographic studies of ubiquitous computing in the wild," in *Proceedings of the 6th conference on Designing Interactive systems*, ser. DIS '06. New York, NY, USA: ACM, 2006, pp. 60–69.

[7] H. Garfinkel, "The origins of the term 'ethnomethodology.'," *Ethnomethodology*, vol. 15, p. 18, 1974.

[8] C. Geertz, "Thick description: Toward an interpretive theory of culture," *Culture: critical concepts in sociology*, pp. 173–196, 1973.

[9] S. Gresham-Lancaster, "The aesthetics and history of the hub: The effects of changing technology on network computer music," *Leonardo Music Journal*, vol. 8, pp. 39–44, 1998.

[10] S. Hewitt, P. Tremblay, S. Freeman, and G. Booth, "HELO: the laptop ensemble as an incubator for individual laptop performance practices," in *Proceedings of the International Computer Music Conference*, 2010, pp. 304–307.

[11] G. Wang, D. Trueman, S. Smallwood, and P. Cook, "The laptop orchestra as classroom," *Computer Music Journal*, vol. 32, pp. 26–37, Mar. 2008.

A SYSTEM FOR MEMORIZING SONGS BY PRESENTING MUSICAL STRUCTURES BASED ON PHRASE SIMILARITY

Yuma Ito
Kobe University

Yoshinari Takegawa
Future University Hakodate

Tsutomu Terada
Kobe University
PRESTO, JST

Masahiko Tsukamoto
Kobe University

ABSTRACT

Players of musical instruments usually memorize musical scores for concerts and live performances. However, memorizing songs requires much effort on a part of the player as they have to play and listen to the song over again. The goal of our study is to construct a system for memorizing musical scores based on the phrase similarity. The proposed system calculates the phrase similarity in the target song, and presents the musical structures and the different points in similar phrases based on the phrase similarity. The learner can understand the musical structure immediately, and can memorize the musical score in a short time because of the reduction of duplicated learning for the similar phrases. Our evaluation results confirmed that our method had advantages compared with conventional musical scores.

1. INTRODUCTION

Players on musical instruments usually memorize musical scores for concerts and live performances. It is important for musicians to memorize musical scores. However, for memorizing musical scores, it needs a great effort on players by playing and listening the song over again. In addition, because it is difficult to memorize musical scores correctly, players sometimes play the same phrase multiple times or forget part of a song due to stress when performing in front of an audience.

On the other hand, a song has musical structures, such as motifs. Figure 1 shows an example of musical structure. There are multiple layers from abstracted layer to detailed layer, as shown in different colors on the figure. The upper area indicates a detailed layer and the lower area indicates an abstracted layer. Additionally, each bracket indicates a phrase belonging to each layer. Most of these musical structures are composed of similar motifs.In addition, information on intersection between motif A and motif A' can be reduced to minimize the learning time. Players can memorize phrases correctly by being aware of the different points among similar phrases.

Therefore, the goal of our study is to construct a system for memorizing musical scores based on the phrase similarity. The proposed system calculates the phrase similarity in a target song, and presents the musical structures and the annotations, such as similarities and differences among similar phrases. In this way, the learner can understand the musical structure immediately, and can memorize the musical score in the short time because of reducing the shared information of the song.

2. RELATED WORK

2.1. Musical Structure representation

ScoreIlluminator[1], whose target users are the same as that of our research, supports the readability of an orchestra score for players and conductors by classifying melodies and accomplishments into groups and assigning colors to the classified objects. However, the purpose of this research is improvement of the readability, while our purpose is memorization of musical scores.

2.2. Support of Memorizing Musical Score

There are a few works that evaluate effective methods to memorize musical scores scientifically. Most of these methods for memorizing musical scores are based on theories derived from the experience of professional players. For example, Bernstein [2] insists that non-conscious memory is weak and may fail to provide the answer to a simple question like "what was the next note?". Therefore the backup memory based on conscious memory, when players are conscious of a musical structure as shown in Figure 1 and recognize the modulation point, is important. Accordingly, how to aid learners to commit musical structure to their conscious memory is a key point in the design of our proposed system.

Synder [3] describes the process of memorizing a song, which is investigated from the view points of cognitive psychology and information theory. Specifically, it is necessary for players to analyze the musical structure, and it is especially important for them to be conscious of the different points of two similar phrases. In our research, we develop a system that prevents the learner from memorizing ambiguously by presenting different points among similar phrases on musical scores, thus helping the learner to memorize musical scores correctly.

3. DESIGN

As mentioned in Section 2, it is important to be conscious of the musical structure of a musical score and the different points among similar phrases. Musical structure is composed of multiple musical layers from abstracted layer to detailed layer as shown in Figure 1. The learner requires the abstracted layers to check brief overview of a song. In contrast, the learner requires the detailed layers to check the momentary states of a song, which are rhythm, pitch, and musical techniques. Because the necessary layer is dependent on the ability or the interest of

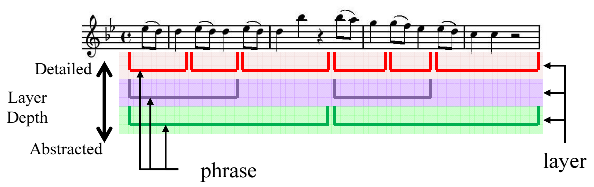

Figure 1. Musical structure

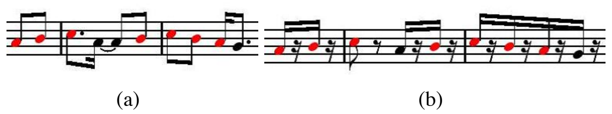

Figure 2. An example of timing

the learner, the proposed system must be able to switch flexibly among presentation of the different layers.

3.1. Phrase Similarity

In this research, the degree of similarity is calculated in units of phrase. The phrases and their hierarchical structure are generated using the method proposed by Hamanaka et al.[4] on the basis of GTTM proposed by Lerdahl et al.[5]. Additionally, the degree of similarity is calculated based on two types of similarities. One is physical similarity, such as fingering, is dependent on musical instruments, and the other is musical similarity, such as pitch and duration, is independent of musical instruments.

3.1.1. Musical similarity

The musical similarity is calculated based on feature values such as pitch and rhythm, which are extracted from each phrase. We employ three feature values; timing (the onset time of each note), pitch (the absolute pitch), and interval of pitch (the difference in pitch between the current note and the one that precedes).

Figure 2 shows an example of timing. We define the onset time of each note on the basis of the first note of a phrase as timing. For example, the two phrases shown in the figure have notes of different durations, but the onset timing is similar because in both phrases there are many notes which have the same onset timing in, as shown by the red notes in the figure.

Our system uses DTW (Dynamic Time Warping) [6] to measure the similarity of two phrases. DTW can be used to measure the similarity between two sequences, each of which may be differently stretched or compressed in time.

The value of musical distance $d_m(i,j)$ is defined by the following equation. N is the number of phrases in a musical layer.

$$d_m(i,j) = w_\mathrm{t} d_\mathrm{t}(i,j) + w_\mathrm{p} d_\mathrm{p}(i,j) + w_\mathrm{i} d_\mathrm{i}(i,j)$$
$$(i = 1,\ldots,N)(j = 1,\ldots,N)$$

Note that $d_\mathrm{t}(i,j)$, $d_\mathrm{p}(i,j)$, and $d_\mathrm{i}(i,j)$ are the distances in timing, pitch, and interval of pitch, which are calculated as a result of DTW. Additionally, w_t, w_p, and w_i are the weighting factor of each element.

Two phrases of which $d_m(i,j)$ is lower than the threshold are defined as musically similar phrases.

3.1.2. Physical similarity

Fingering in playing a musical instrument is an example of physical similarity. Because position and physical information, such as fingering, is almost mechanical memory, it is important to back this up with conscious memory by presenting the fingering similarity. Additionally, fingering is different for each musical instrument. In other words, physical information such as fingering is dependent on the musical instrument.

Physical similarity is calculated by DTW in the same way as musical similarity. The physical distance between i-th and j-th is defined as $d_p(i,j)$. Strings, frets, and musical performance techniques such as arpeggio and hammer-on are an element of the DTW of physical similarity. $d_p(i,j)$ is the sum total of the results of DTW for each element. Two phrases of which the $d_p(i,j)$ is lower than the threshold are defined as physically similar phrases.

3.2. Method of Memorization

We explain a proposed method for memorizing scores with Figure 3 and Figure 4. The proposed system has two modes: *All phrases presentation mode* and *Similar phrases presentation mode*. The learner uses the *All phrases presentation mode* to learn the structure of the song or to select a base phrase that is used in the *Similar phrases presentation mode*. Additionally, users learn the similarities and differences between the selected base phrase and other phrases with the *Similar phrases presentation mode*.

3.2.1. All phrases presentation mode

Figure 3 shows an example of the *All phrases presentation mode*. Each rectangle in the figure denotes a phrase. In our system, the learner can change the presented layer in these layers freely. The learner uses this mode to learn the structure of the current layer or to select a base phrase used in the *Similar phrases presentation mode*.

3.2.2. Similar phrases presentation mode

This mode presents a base phrase selected by the learner in the *All phrases presentation mode* and the phrases that are similar to. We propose two types of content presentation as shown in Figure 4. The left-hand diagram shows a general musical score, and the right-hand diagram shows a summary of the similar phrases. The rightmost scores are guitar tabs. Details of the example in Figure 4 are given below, and the Roman numerals in the black dots in Figure 4 correspond to the following list.

(i) The phrases surrounded by a solid rectangle are base phrases, whereas, similar phrases are surrounded by a dotted rectangle. The numbers next to the rectangles, show the degree of similarity. In the right-hand diagram of Figure 4, the base phrase is placed at the top of the list, and other similar phrases are arranged in order from the highest degree of similarity to the lowest. The learner can easily understand the similar phrases and their location in the musical score.

(ii) The number of identical phrases is indicated with a numerical value, such as "×2", which appears next to the base phrase. In this way, duplicated information is reduced because the learner does not have to re-memorize the phrase when it reappears later on in the song.

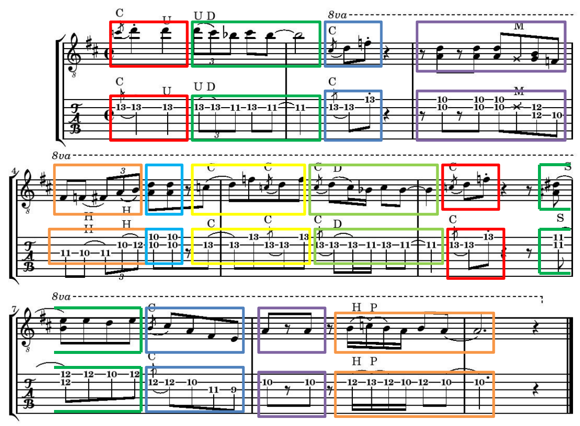

Figure 3. An example of All phrases presentation mode

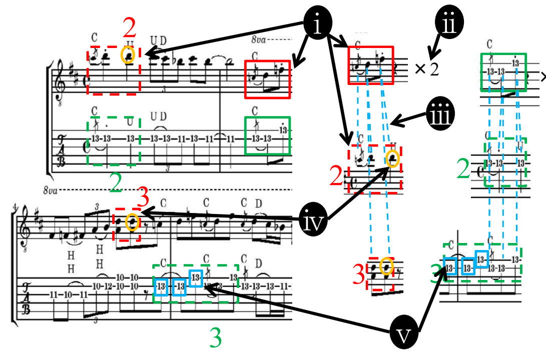

Figure 4. An example of the similar phrases presentation mode

(iii) When the timing or fingering of a note in a similar phrase is the same as that of the base phrase, they are connected by a dashed line. The learner can understand the similarity of timing or fingering.

(iv) When there are the notes in the similar phrases that are same as those of the base phrase in regard to pitch, timing, and fingering, they are surrounded by circles. The learner can understand the similarity of each feature value.

(v) When there are notes in the similar phrases that are different from those of the base phrase, they are surrounded by squares. Using this information, the learner can understand the differences relating to each feature value, and is prevented from incorrect memorization.

4. IMPLEMENTATION AND EVALUATION

We implemented a prototype system for memorizing musical scores. The prototype stores musical score including meta-data such as pitch, timing, and fingering, in XML. It calculates the degree of the similarity among phrases based on the musical score data. We implemented the prototype using Microsoft Visual C# 2008 on Windows 7.

We evaluate the proposed method by comparing with a conventional method.

Table 1. Memorization time

	Song A	Song B	Song C	Song D	ability value
Sub. 1	11	21	19	16	1.00
Sub. 2	16	20	20	25	1.03
Sub. 3	25	19	26	22	1.61
Sub. 4	25	35	27	40	2.91

Table 2. Ratio of memorization time

	Song A	Song B	Song C	Song D	average	S.D.	p-value
Sub. 1	0.71 0.80	1.78 1.74			1.26	0.5	0.22
Sub. 2		1.64 1.62	1.02 2.09		1.59	0.38	0.04
Sub. 3			0.85 1.74	0.86 0.56	1.00	0.44	0.50
Sub. 4	0.55 0.63			0.86 0.57	0.65	0.57	0.01
average	0.67	1.69	1.43	0.71			
S.D.	0.09	0.07	0.51	0.15			
p-value	0.004	0.0002	0.12	0.02			

In the evaluation, the subjects practiced the song until they had memorized it using the two methods: the conventional method, which uses conventional musical scores, and the proposed method, which uses the musical scores created by the prototype. We measured how long subjects took to memorize each trial score. Each test subject memorized four; two songs with the conventional method and two songs with the proposed method. The subjects were permitted to use the electric guitar and listen to the songs during the evaluation as needed. Furthermore, before the subjects worked on the task with the proposed method, we instructed them on how to operate the prototype and how to learn songs with it.

All the songs we used belong to the rock music genre and they are composed of approximately 6 to 9 bars. We call these songs *Song A*, *Song B*, *Song C* and *Song D* for convenience.

There were four subjects, all of whom were university students in their early 20s who major in engineering, can read musical scores, and play the electric guitar as well. The subjects had never listened to or practiced the trial songs before.

Table 1 shows the results of the experiment. Each value in Table 1 denotes the time in minutes that a subject took to memorize a song, which is referred to as memorization time. The gray cells indicate the memorization time applied to the proposed method, and the white cells indicate the memorization time applied to the conventional method.

Each subject had different level of ability for reading musical scores and different musical techniques. Each song presented different difficulties. Therefore, we normalize the memorization time based on the memorization time applied to the conventional method. Specifically, we define the ability of each subject as follows:

$$a_n (n = 1,2,3,4)$$

Note that "n" corresponds to the number assigned to each subject. We determine the a_n based on the memorization time of the conventional method. For example, since Subject 1 memorized Song C and Song D using the conventional method, we can compare Subject 1 with Subject 2 and Subject 4 who also memorized Song C or Song D using the conventional method. The equation is as follows:

$$a_1 = \frac{16}{25}a_2 = \frac{19}{27}a_4$$

We formulate the equation for each subject, and define the average number of multiple solutions of a_n as the final a_n, as shown in the rightmost in the table.

Furthermore, we calculate the primary memorization time in the case in which the subject did not use the proposed method in contrast to the memorization time with the proposed method. The primary memorization time is determined on the basis that the ratio of the primary memorization time to the time using the conventional method is equal to the ratio of abilities. For example, if we compare Subject 1 with Subject 2, the primary memorization time x is determined as 15.6 by following equation.

$$\frac{x}{16} = \frac{1.00}{1.03}$$
$$x = 15.6$$

This value indicates that the primary memorization time of Subject 1 was reduced from 15.6 min to 11 min by using the proposed method.

We determine the ratio of the memorization time of the proposed method to primary memorization time, and we refer to this as the ratio of memorization time. The ratio of memorization time is determined by the following equation.

$$\frac{11}{15.6} = 0.71$$

This value indicates that the memorization time is reduced by approximately 30%.

Table 2 shows the ratio of memorization time in all the combinations. The three columns on the right-hand side and the bottom three rows of Table 2 show the value of average, standard deviation, and p-value that indicates the statistical significance by t-test for each subject and each song.

The memorization time of Song A and Song D are reduced. The proposed system worked effectively on these songs. The reason is that over half of the phrases are highly similar to each other. On the other hand, the memorization time of Song B and Song C are increased. These songs have some complex rhythms that take time to master and accordingly, the memorization time was increased.

Subject 4 was able to reduce the memorization time of all the songs he memorized with the proposed method, and it had significance. He stated that the ranking of similarity was useful because he could select the phrases in order of the ease of memorization. In contrast, Subject 2 was not able to reduce the memorization time, and it had a significance. According to his comments, the prototype system had a lot of functions and he could not understand them all. For this reason, he could not get useful information from the prototype system to aid his memorization.

5. CONCLUSIONS

We constructed a system for memorizing songs by presenting musical structures based on phrase similarity. It presents the musical structure of the song, similar phrases, and differences among similar phrases. Form the result of evaluation, although the effectiveness of the proposed system depends on the musical ability of the subject and characteristics of the song, the memorization time was reduced by using the proposed system.

Future work will include experiments using subjects of varying ability and more extensive experiments. Additionally, we intend to apply the proposed system to instruments other than the guitar.

6. ACKNOWLEDGEMENTS

This research was supported in part by a Grant in aid for Precursory Research for Embryonic Science and Technology (PRESTO) from the Japan Science and Technology Agency and by a Grant-in-Aid for Scientific Research (A)(20240009) and Scientific Research for Young Scientists (b)(21700198) from the Japanese Ministry of Education, Culture, Sports, Science and Technology.

7. REFERENCES

[1] M. Matsubara, H. Okamoto, T. Sano, H. Suzuki, S. Nobezawa, and H. Saito: Scoreilluminator: Automatic Illumination of Orchestra Scores for Readability Improvement, *The Proceedings of the 2009 International Computer Music Conference (ICMC2009)*, pp. 113–116 (2009).

[2] S. Bernstein: With Your Own Two Hands, *G. Schirmer, Inc*. (1981).

[3] R. Snyder: Music and Memory, *The MIT Press* (2001).

[4] M. Hamanaka, K. Hirata, and S. Tojo: Implementing "A Generative Theory of Tonal Music", *Journal of New Music Research (JNMR)*, Vol. 35, No. 4, pp. 249–277 (2007).

[5] F. Lerdahl and R. Jackendolf: A Generative Theory of Tonal Music, *The MIT Press* (1983).

[6] R. Bellman: Dynamic Programming, *Princeton University Press* (1957).

_ICMC2012 NON-COCHLEAR SOUND
LJUBLJANA _9.–14. SEPTEMBER

A WIRELESS, REAL-TIME CONTROLLER FOR THE ENSEMBLE AUDIO-VISUAL SYSTEM.

Anton Jidkov
Music Technology
Computing and Engineering
University of Huddersfield
anton.jidkov@gmail.com

Dr Ian Gibson
Music Technology
Computing and Engineering
University of Huddersfield
i.s.gibson@hud.ac.uk

Mark Hildred
Apollo Creative
mark@apollocreative.co.uk

ABSTRACT

This research investigates development of sound diffusion software for the Apple iPhone and iPad. It is specifically written for Apollo Creative's 'Ensemble' system, which is an interactive audio, lighting and video software and hardware package. It allows users to design software instruments for sound installations and shows, reacting to a range of input sensors. The iPhone software controls the surround sound parameters of Ensemble. The user interface is designed to accommodate a range of users, including educators, composers and sound designers. The software incorporates the OpenFrameworks library [9] and communicates with 'Ensemble' using the Open Sound Control (OSC) [15] protocol over wireless TCP/IP.

1. THE ENSEMBLE SYSTEM

The Apollo Ensemble [1] is a system designed for teachers and special needs specialists, allowing them to configure interactive sensory environments for individuals with a range of disabilities. Ensemble also has applications in the areas of exhibitions, artistic installations and children's play areas.

Ensemble is a switch operated system that is an evolution of Midicreator [7]. Switch operated hardware allows people with disabilities to have a range of control methods that can be mapped to output stimuli via specialised software. Similar examples include the Skoog [11], a tactile music instrument; MidiMate [5], an access device for electronic keyboards and Quintet [13].

The Ensemble system is split into three main components: a range of input sensors for detecting movements; designer and player software running on a PC; and output devices which can include sound, lighting, video, image and sensory equipment.

1.1 The Hub interface

The Ensemble Hub, figure 1, forms the main USB interface for the PC, featuring four sockets for simple on/off switches and two for variable sensors. It also contains a 433MHz transmitter for controlling proprietary sensory equipment and a 2.4GHz module for wireless sensors. A basic portable setup can be achieved by using the Hub together with a laptop or netbook PC.

Figure 1. Ensemble Hub

1.2 Input sensors

Due to the nature of the special needs market, there are a large number of standard 'assistive technology' switches available for a wide variety of needs. Past system development has focused on wireless adaptors. It is important that there is a clear link between cause (i.e. sensors employed) and effect (sonic output) in special needs work [2]. Non-contact sensors can be problematic in these environments. Also sensors need to be robust and simple to use.

The current range of sensors include:

- Connect – adaptor to allow up to four switches to be linked wirelessly to the Ensemble Hub.
- Dice – each side of the sensor can be used to trigger a different sound or effect.

Figure 2. Players using the dice sensor within a foam cube.

- Press – sensitive pressure pad that produces a variable signal depending on the amount of pressure applied.
- Tilt – produces a variable pitch and roll output as it is rotated.
- Squeeze – an air pressure bulb whose output varies with the pressure applied.
- Dual – adaptor to allow up to two variable sensors to be linked wirelessly to the Ensemble Hub.