Development of a Speech-Driven Embodied Laser Pointer with a Visualized Response Equivalent to Nodding

Hiroyuki Nagai[*1], Tomio Watanabe[*1,*2] and Michiya Yamamoto[*1]
∗1 Okayama Prefectural University
111 Kuboki, Soja, Okayama, 719-1197, Japan
∗2 CREST of Japan Science and Technology Agency (JST)

Abstract

InterActor is the speech-driven embodied entrainment character for activating human interaction and communication by embodied rhythms between speech and body motions such as nodding in human face-to-face conversation. In this paper, a new speech-driven embodied laser pointer InterPointer is developed for supporting embodied interaction and communication for both a lecturer and audiences. InterPointer is a pointing device in which the beam pattern is changed from a point to an oval with a visualized response equivalent to nodding in the same timing as InterActor to speech input.

1 Introduction

According to the rapid progress of information and communication technology, it has become possible easily to communicate with distant people in everyday life. In human face-to-face communication, not only verbal message but also nonverbal behavior such as nodding and body motions are rhythmically related and mutually synchronized between talkers [1]. It is to be desired that people can share embodied rhythm by the entrainment between human speech and nonverbal actions and motions in remote communication.

We have developed the speech-driven embodied interactive actor called InterActor which has both functions of speaker and listener by generating expressive actions and motions coherently related to speech input. We demonstrated that the system can be effective for supporting human interaction and communication between remote talkers [2].

It is difficult for a lecturer on the platform to see audiences' reaction because they are separated in a conference room, and they are not entrained. Therefore, it is difficult for the lecturer to present the lecture like face-to-face communication. It is expected to support interactive communication by introducing the synchrony of embodied rhythms.

In this paper, a new speech-driven embodied laser pointer InterPointer is developed for supporting embodied interaction and communication in a lecture. InterPointer is a pointing device in which the beam pattern is changed from a point to an oval with a visualized response equivalent to nodding in the same timing as InterActor to speech input.

2 Development of a speech-driven embodied laser pointer

2.1 Concept of InterPointer

Figure 1 shows the concept of a speech-driven embodied laser pointer "InterPointer". InterPointer is the system for supporting interaction with lecturer and audiences through visualized response equivalent to nodding of listener on the basis of speaker's speech. Even if audiences' reaction cannot be seen and it is hard to talk on the platform, a lecturer can make presentation with ease by using InterPointer that can feed back a typical nodding reaction. Audiences can also

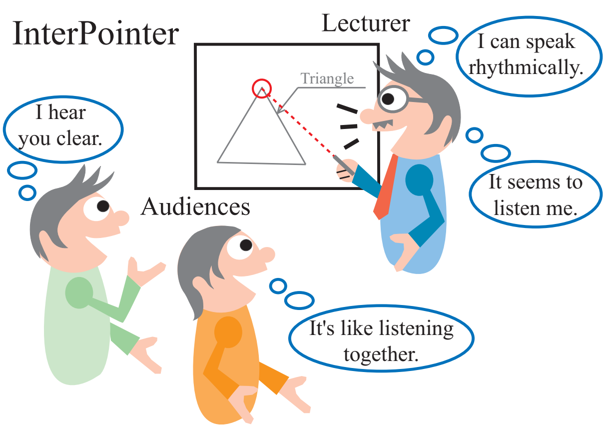

Figure 1: Concept of InterPointer.

enjoy good presentation, because they can share embodied rhythm with the lecturer and feel a sense of togetherness by seeing InterPointer's nodding reaction.

2.2 Outline of InterPointer

Figure 2 shows the outline of InterPointer. InterPointer utilizes the change of beam patterns of a laser pointer LP-2000 (CABIN) for the visualization of nodding. LP-2000 can generate four kinds of beam patterns by laser light reflection in the mirror which is vibrated by two electromagnetic vibrators (Figure 3). At the same time, slight vibration arises at the timing of pattern generation.

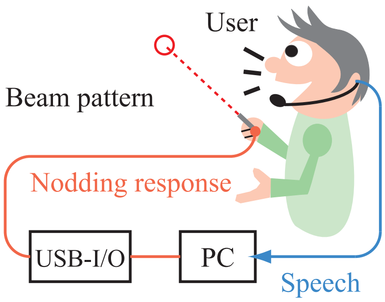

Figure 2: Outline of InterPointer.

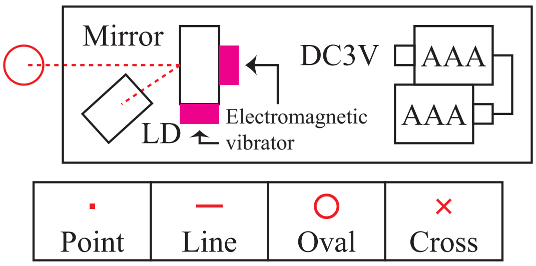

Figure 3: Outline of LP-2000.

Users can feel the nodding response of listener by change of beam patterns. Response time of a nodding was set to 200 ms. InterPointer is controlled by the USB-I/O from PC equipped with the microphone input and the USB connector. Table 1 shows beam patterns of InterPointer.

2.3 Development of InterPointer

Figure 4 shows InterPointer. InterPointer has installed the microcomputer PIC16FL84A (MICROCHIP) which can set up input and output for pattern generation. The microcomputer PIC16FL84A

Table 1: Beam patterns of InterPointer.

	USB-I/O cable is not connected.		Nodding response		
			OFF	ON	
Pattern	·	—	·	·	—
	○	×		○	×

was connected to the button input part and it generated pattern switch signals. InterPointer had a connector in order to connect an USB-I/O cable and to receive a nodding signal from PC.

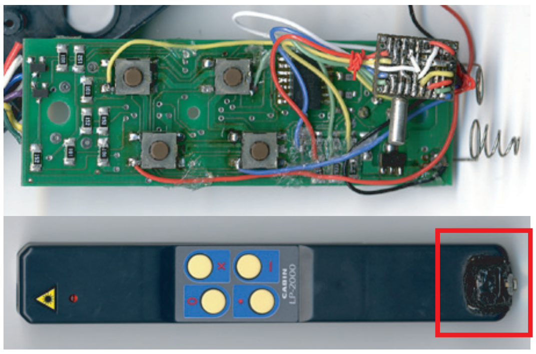

Figure 4: InterPointer.

Original cable called USB-I/O cable was developed to connect USB with InterPointer as shown in Figure 5. USB-IO was built in the USB-I/O cable. The control chip of USB-I/O (Morphy Planning) was a CY7C63001A-PC (Cypress Semiconductor). It supported USB 1.1 (Low-speed, 1.5 Mbps) and was bus-powered.

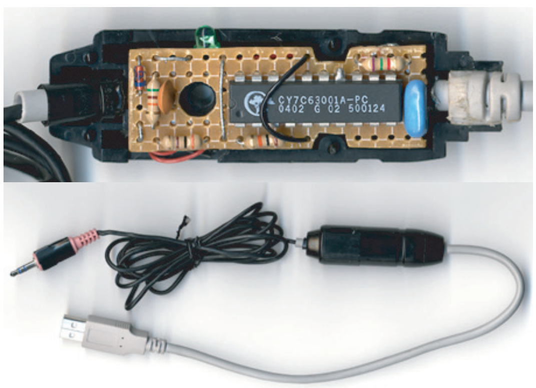

Figure 5: USB-I/O cable of InterPointer.

2.4 Listener's interaction model

InterPointer is the laser pointer that can change beam patterns in the equivalent timing to listeners nodding, which is one of the typical patterns of non-verbal interaction, and it plays an important role in smooth interaction in human communication. As for the prediction model of nodding on human face-to-face communication, the MA (Moving-Average) model which estimated the nodding $y(i)$ as the weighted sum of the binary speech signal $x(i)$ in each video frame of 1/30 second was introduced as follows [3].

$$\hat{y}(i) = \sum_{j=1}^{J} a(j)\, x(i-j) + w(i) \qquad (1)$$

$a(j)$: linear prediction coefficient, $w(i)$: noise

3 Evaluation of the system

3.1 Experimental method

The effectiveness of the InterPointer system was examined by sensory evaluation. Figure 6 shows the experimental setup of the system. One of paired subjects was a lecturer who makes a presentation in front of a screen, and the other was an audience for the presentation. The following four modes from (a) to (d) were put to paired comparison in the experiment.

(a) Point (the beam pattern was a point, and doesn't change during the experiment)
(b) Oval (the beam pattern was an oval, and doesn't change during the experiment)
(c) Model (the beam pattern changes from Point to Oval in the nodding timing by MA-model and returns to Point after 200 ms)
(d) Exponential (the beam pattern changes from Point to Oval in the timing of exponential distribution and returns to Point after 200 ms)

The subjects were instructed to experience a presentation and choose better mode out of randomly selected two modes from (a) to (d). The exponential distribution was generated as follows.

$$T = -3\ln(1-\lambda), \quad \lambda = [0, 1] = (0.1 \sim 0.9) \qquad (2)$$

The average of exponential distribution was set at 3 seconds which denotes the mean nodding interval.

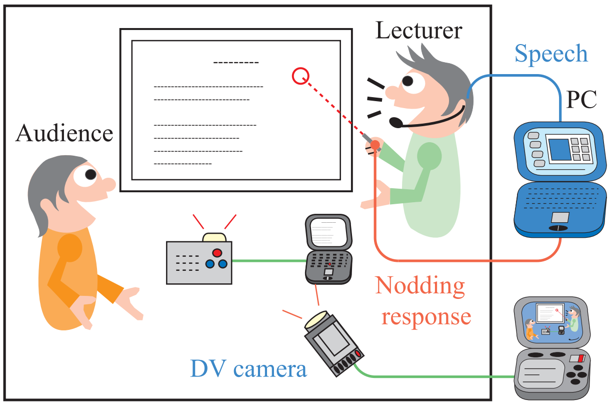

Figure 6: Experimental setup of InterPointer.

The experiment was examined in the following steps :

Steps 1 ; Subjects tried InterPointer, so that they can get used to the system.
Steps 2 ; They were instructed to divide into a lecturer and an audience to perform paired comparison by one-minute use for each mode.
Steps 3 ; A lecturer and an audience swapped roles in the same way.

Figure 7 shows the experimental scene of InterPointer and Figure 8 shows the example of change of beam pattern (200 ms). The contents of slides were the formulas of area, such as a triangle and a trapezoid, as subjects can explain easily and straightforwardly. The experiment was examined by 10 pairs of 20 Japanese students. The outline of a system was not explained to subjects beforehand.

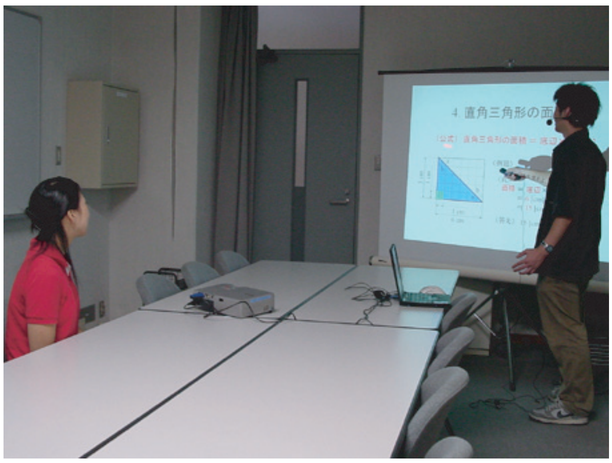

Figure 7: Experimental scene of InterPointer.

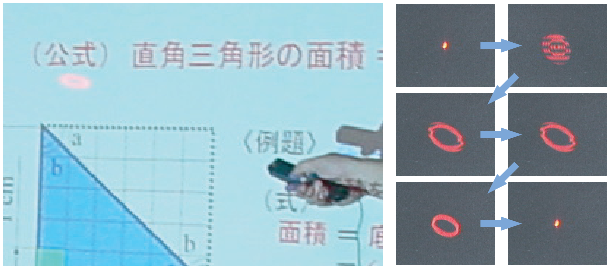

Figure 8: Example of change of beam pattern (200 ms).

3.2 Experimental result

Table 2 shows the result of paired comparison for the lecturer and audience. The number in the table shows that of subjects who preferred the line mode to the row mode. Namely, this number is the number of the subjects who answered that it was more desirable in the mode of each line. The Bradley-Terry model was assumed to evaluate the preference of mode by InterPointer quantitatively, defined as follows [4].

$$P_{ij} = \frac{\pi_i}{\pi_i + \pi_j} \tag{3}$$

$$\sum_i \pi_i = const. (= 30) \tag{4}$$

π_i : intensity of i,
P_{ij} : probability of judgment that i is better than j

Here, π_i shows the intensity of preference of mode by InterPointer. The model enables to determine the preference based on the paired comparison. The Bradley-Terry model assumed by using the result of paired comparison is also shown in Table 2. To approve the matching of the model, the goodness-of-fit test and likelihood rate test were applied to this Bradley-Terry model. As a result, the model was not rejected and π was validated.

By the result of paired comparison, the proposed model was highest. This indicates that the change from Point to Oval by InterPointer is preferred to Point by both lecturer and audiences. Much affirmative opinions for InterPointer were also seen by the comments after the experiment.

Table 2: Result of paired comparison of InterPointer.

IP	(a)	(b)	(c)	(d)	Σ	π
(a)		23	21	23	67	8.54
(b)	17		12	8	37	3.87
(c)	19	28		24	71	9.48
(d)	17	32	16		65	8.11

(a) ▪ (Point) (c) ▪ →○ (Model)
(b) ○ (Oval) (d) ▪ →○ (Exponential)

4 Conclusion

In this paper, a new speech-driven embodied laser pointer was proposed, and the prototype of the system InterPointer was developed for supporting embodied interaction and communication for both a lecturer and audiences. The effectiveness of InterPointer was demonstrated by sensory evaluation.

Acknowledgements

This work under our project E-COSMIC has been supported by CREST (Core Research for Evolutional Science and Technology) of JST (Japan Science and Technology Agency).

References

[1] N. Kobayashi, T. Ishii, T. Watanabe, “Quantitative Evaluation of Infant Behavior and Mother-Infant Interaction”, *Early Development and Parenting*, pp. 23-31, 1992.

[2] T. Watanabe, “E-COSMIC : Embodied Communication System for Mind Connection”, *Proceedings of the 13th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 2004)*, pp. 1-6, 2004.

[3] T. Watanabe, M. Okubo, M. Nakashige, R. Danbara, “InterActor : Speech-Driven Embodied Interactive Actor”, *International Journal of Human-Computer Interaction*, Vol. 17, No 1, pp. 43-60, 2004.

[4] R. A. Bradley and M. E. Terry, “Rank analysis of incomplete block designs”, *Biometrika*, Vol. 39, pp. 324-345, 1952.