Rep. Fac. Sci. Engrg.
Saga Univ.
36-1 (2007),39-44

Reports of the Faculty of Science and Engineering,
Saga University, Vol. 36, No.1, 2007

e-learning system with a confidence evaluation through voice response analysis

By
Kohei Arai* and Hiroshi Yoshida**

Abstract: An e-learning system with a confidence evaluation through voice response analysis is proposed. The system does care about the confidence level of the students when the students respond to the questions raised from the systems through human voice analysis. Otherwise, the students et a trouble as they pass step-by-step based exercises without confidence. So that the proposed system evaluates their confidence during exercises with human voice analysis, how do they feel. Through experiments with 10 of students, it is found that the proposed system achieved 10% improvement of the final achievement test result in comparison to the system without confidence evaluation.

Key words: e-learning, voice analysis, confidence level evaluation, pitch and power, clustering

1. Introduction

Advances in technology and the growth of e-learning to provide educators and trainers with unique opportunities to enhance learning and teaching in corporate, government, healthcare, and higher education is covered in the International Journal on E-Learning: IJEL[(1)].

e-learning is popular in the world. Fundamental system is developed from Web Based Training: WBT in the early 1980s. National Technological University utilizing WBT as well as e-Learning systems without any campus was established in Colorado, U.S.A. in 1984[(2)]. Meanwhile, virtual university was initiated in European countries in 1987 and expanded to the European Community Action Scheme for the Mobility of University Students: ERASMUS[(3)]. On the other hand, virtual university in Asia so called University Mobility in Asia and the Pacific: UMAP was started in 1998[(4)]. There are two types of e-learning, on-demand and live basis e-learning systems and also are two major standard systems, Shareable Content Object Reference Model: SCORM[(5)] and ISO/IEC JTC1 SC36 of Information Technology for Learning, Education and Training: ITLET[(6)]. Non-Profitable Organization: NPO of e-learning consortium in Japan adopted SCORM. They developed Japanese version of SCORM version 1.3. Advanced Distributed Learning Initiative: ADL which was established for promoting standardization of learning systems in 1997 developed SCORM standard for interoperable learning management system and digital contents[(7)]. e-learning system consists digital contents, execute engine and software tools. There are so many types of media assets of the digital contents such as 2D, 3D graphics, audio and video. As for the authoring tools, HTML, XML tags, FLASH of time line, Autgorware for eyecon and Toolbook, Director of hyperlink are widely used. There also are so many content scripters of Java, Javascript, ActionScript, XSL, and Simulations. Meanwhile, e-learning content XML which consists of content and semantic structure with XSL and metadata are available for the data structure. Web browser, XML, Java and Plug-in of FLASH, PDF SVG as well as RealOne and WMP of streaming are well known as the execute engine.

Using the existing and widely used e-learning system, students, trainees etc. precede a step-by-step learning processes getting through a mini-test without confirmation of confidence level of the students. Due to the fact that the mini-test is conducted by choosing a candidate of the answers to the questions from the several candidates, the students may precede one step further if the chosen candidate is correct and even if the students feel any confidence at all. Then the students are used to get a trouble when they finish the lessons[(8)]. The students do not understand which steps are not good at. The proposed e-learning system has such a capability of confirmation of a students' confidence level during the students take a mini-test with their voice. There are some methods for confirmation of confidence level. The most typical and economic way is voice recognition. Microphones are cheaper than cameras. The accuracy of voice and face recognition are almost same so that the proposed e-learning system utilizes voice recognition. When they answer to the question in the mini-test with

Received on 28 Apr.2007
*Department of Information Science
**Funai Electric Co., Ltd.

voice, the proposed system recognizes they confidence level. If they select a correct answer with confidence, they can precede one step further. If else, they have to have another mini-test in the same step. Although it takes much longer time when they learn with the proposed system rather than with the conventional systems, the score for the final examination for the proposed is expected to be greater than that for the conventional systems.

2. Proposed e-learning system

2.1 System Configuration

The proposed e-learning system configuration is based on the typical client server system. In principle, e-learning contents are stored in the server side database system and are written in the web browser basis with the PHP Version4.3.8 together with Apache Version 1.3.31. On the other hand, client side has Web Browser of Internet Explore ver.6.0. On the other hand, confidence level evaluation software was developed with the C language. The detailed method description is in the next section.

2.2 Confidence level evaluation with human voice

There are not so small number of methods for human voice recognitions [9],[10]. Pitch frequency and power of voice are selected as representatives of the students' confidence level. Because of the proposed e-learning system needs a tool for only check the confidence level, (a) answer with full confidence or (b) answer without full confidence so that it was thought that simple two features would be enough. The typical relation between two features and emotion is illustrated in Fig.1.

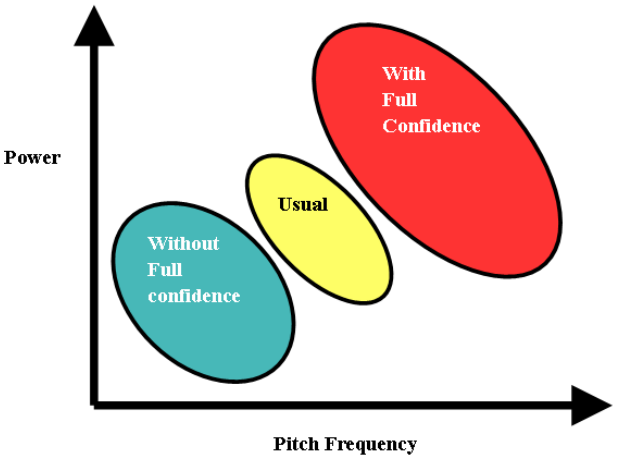


Fig.1 The typical relation between two features, pitch frequency, power of voice and students' confidence level, answer with/without full confidence

Pitch frequency is high and power of voice is also high implies that the students answer with full confidence while both pitch frequency and power of voice are low means that the students answer without full confidence.

The pitch frequency is defined using autocorrelation function.

$$R = \sum X(t)X(t-\tau) \quad (1)$$

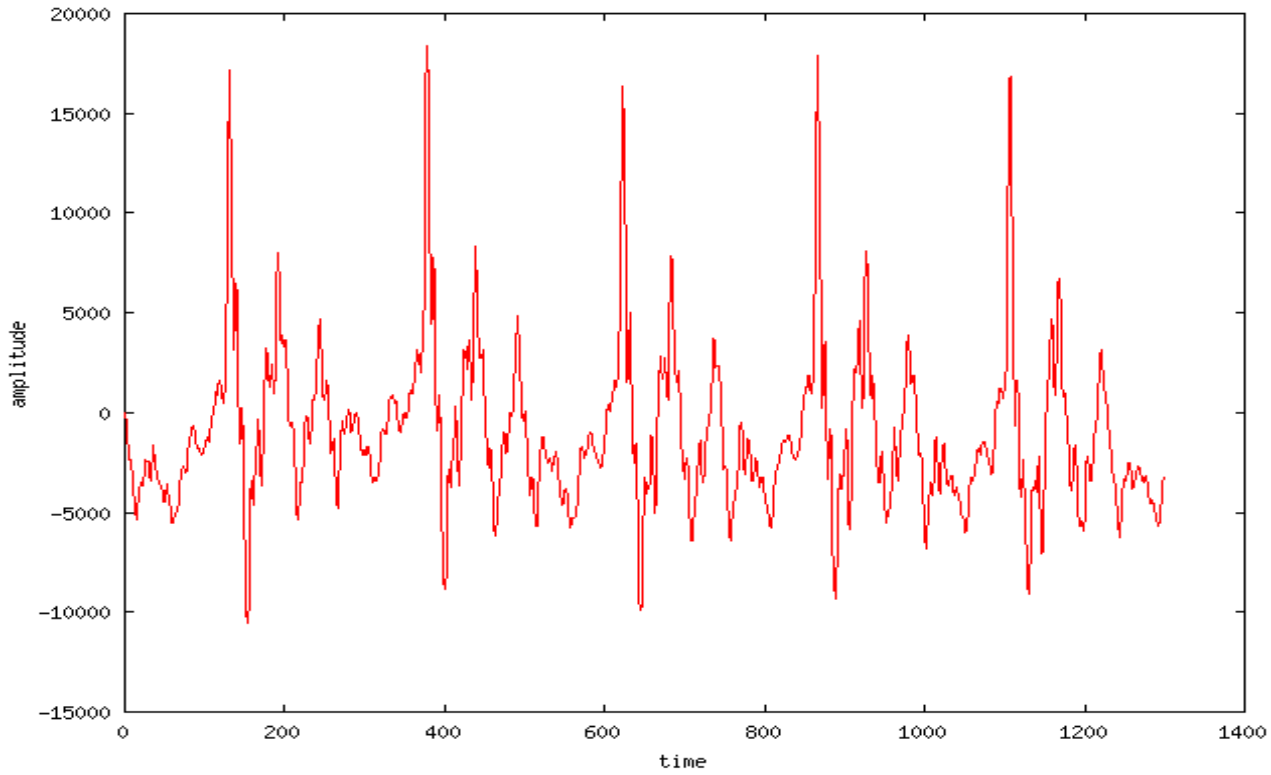
The typical waveform of human voice and its autocorrelation are shown in Fig.2. The pitch frequency is the frequency at the peak of the autocorrelation.
On the other hand, power of human voice is defined as a root mean square of waveform of human voice,

$$P = \sqrt{X^2(t)} \quad (2)$$

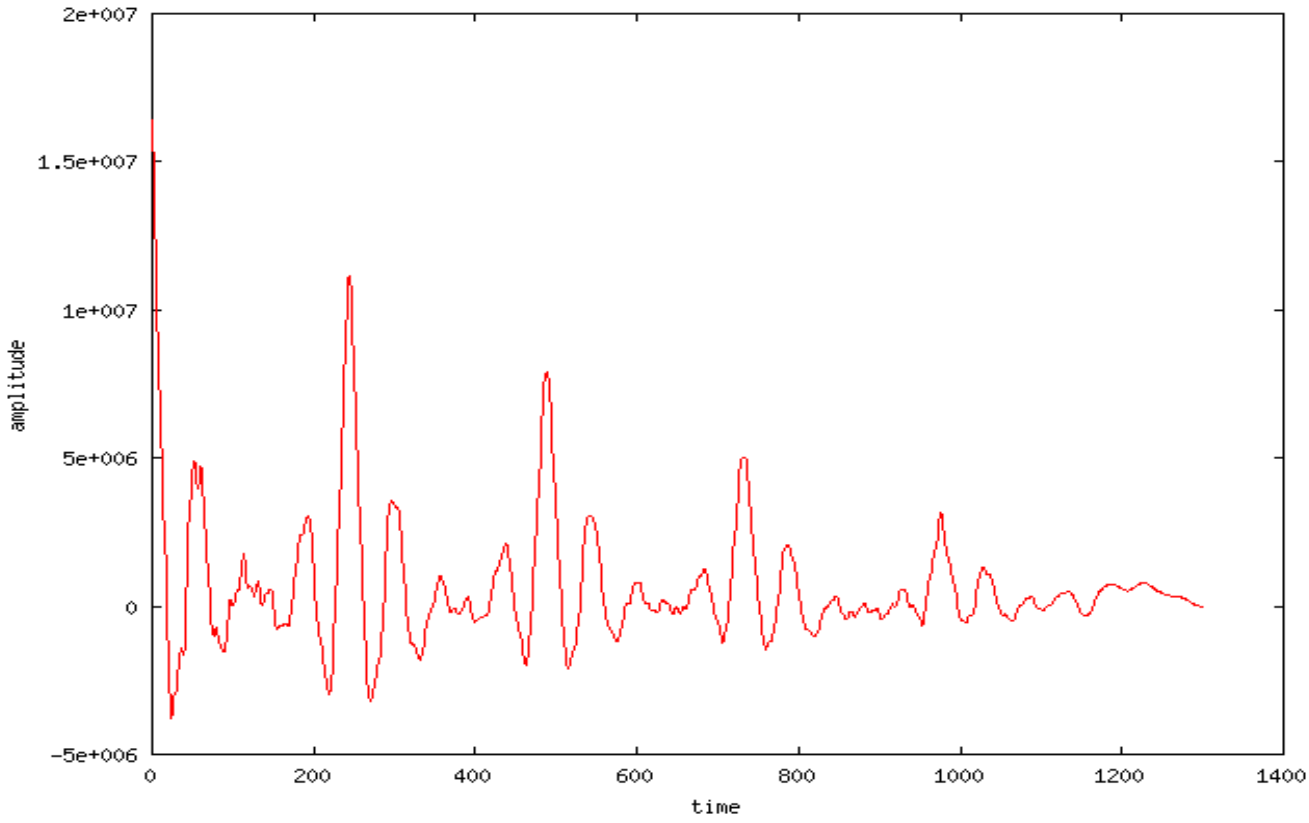
Two features, pitch frequency and power of human voice create a feature space. A human voice is expressed as a data point in the feature space. Next thing to do is to classify the data points into two categories, with and without full confidence. In order to separate students' answer to two categories, with and without full confidence, typical clustering techniques is used. Real time clustering is now widely available which allows classify the data points into several categories, or clusters. For the proposed system, dendrogram, based clustering is used. Closest data points are merged firstly. Then the second nearest data point or cluster center are merged together. These merging processes are contiguously continued and are finished at the number of clusters reached previously designated number of clusters. The definition of distance between data points or cluster centers and the rule for merging are based on WARD method as is well known as one of hierarchical clustering methods.

The typical distribution of data point is illustrated in Fig. 4. In the Fig.4, The data points marked with red ellipsoid are that students answer with full confidence while those marked with green ellipsoid (these data points are out of range of the pitch frequency) are that students answer without full confidence. Blue circle denotes noise. The data points are, sometime, suffered from the surrounding noise

First, the proposed system requires student's name in a normal condition. Then the system calculates pitch frequency and power of human voice and plot the data point as normal condition. Second, the system requires an answer to the question raised from the system then calculate pitch frequency and power of voice. After that the system also calculates the angle between the data point of normal condition and the data points in concern. In order to avoid the noise, the system calculates the angle rather than the distance between the data point of normal condition and the data points in concern.
If and only if the system decided that the students answer to the question from the mini-test with full confidence then proceed the next step and if not, the students have to have another mini-test.



(a) The typical waveform of human voice



(b) The autocorrelation of the typical human voice of (a)

Fig.2 The typical waveform of human voice and its autocorrelation

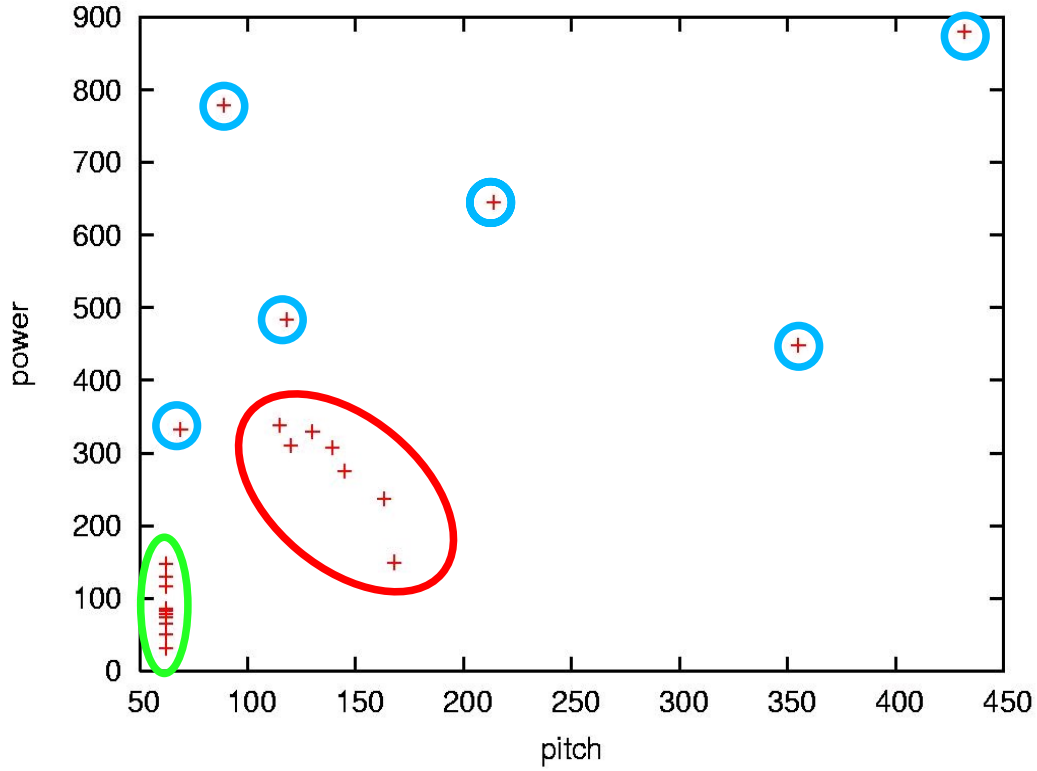


Fig.3 An example of the distribution of data points in the feature space.

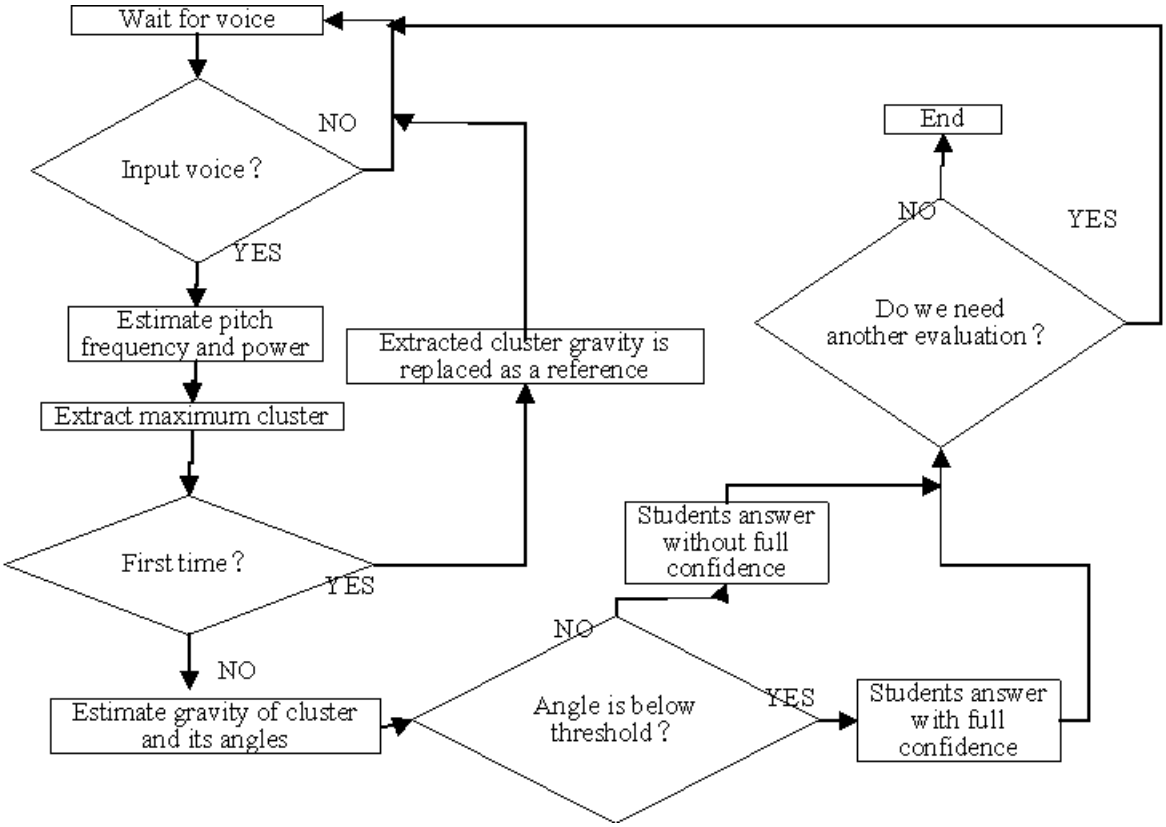


Fig.4 Flow chart of the process for evaluation of confidence level of the students

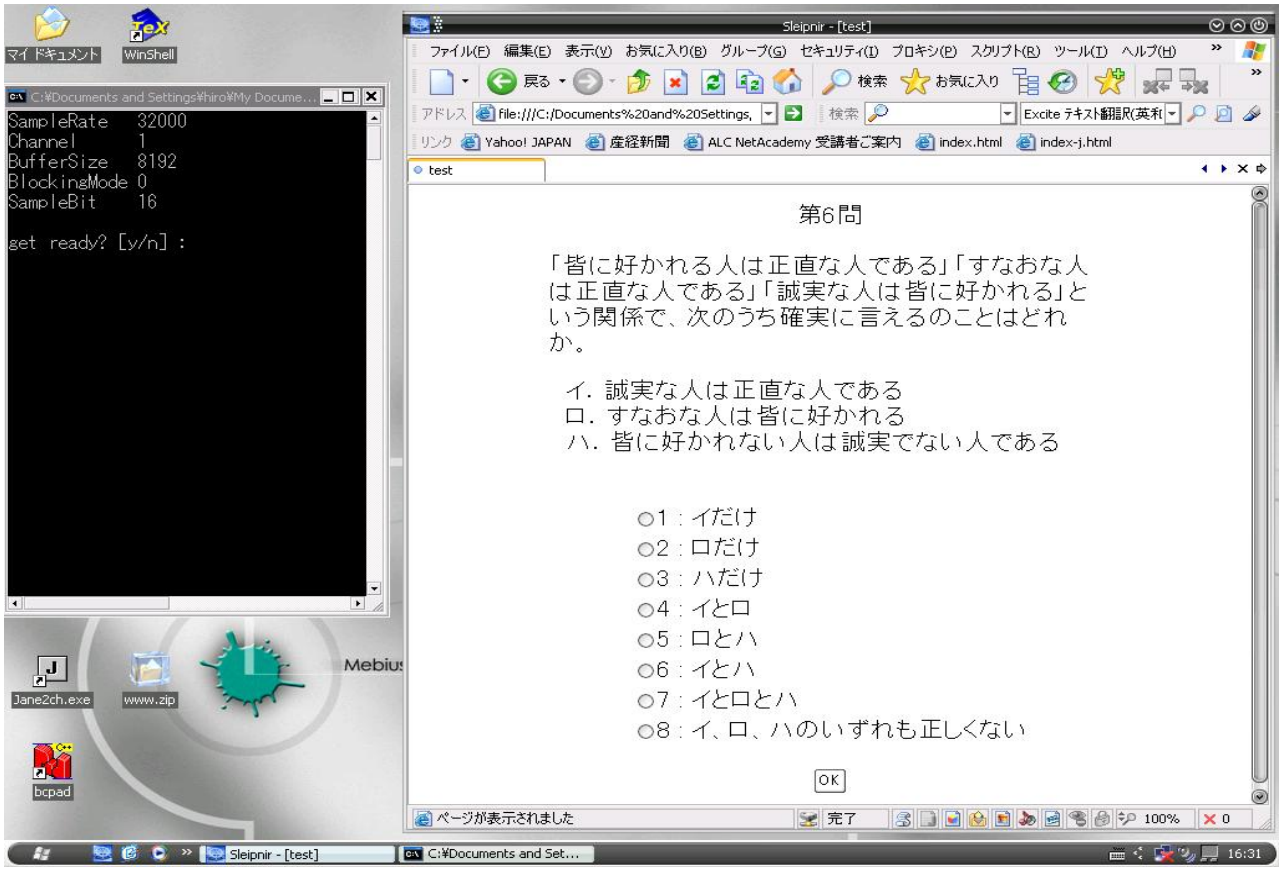


Fig.5 An example of the questions displayed on the web browser of the client side of e-learning system.

Table 1 Score and the time required for the learning with 10 mini-test questions and another 10 final examination test questions extracted from SPI test for the existing and the proposed e-learning system with students' confidence evaluation with human voice recognition.

		Mini-tests	Final examination	Improvement(%)
The e-learning system without confidence evaluation	Score	54.3	51.4	-5.3
	Elapsed time	13'11"	13'03"	
The e-learning system with confidence evaluation	Score	51.4	68.6	33.5
	Elapsed time	29'33"	8'58"	

Process flow for the evaluation of the students' confidence level is illustrated is Fig 4.

4. Experiments

4.1 Procedure

10 of students are selected for experiment of the proposed e-learning system. The questions are selected from the SPI tests without any intention. One of the examples of the questions is shown in Fig.5. All the questions are displayed on the web browser of the client side. First, the students input their name with their voice in a normal condition. Then the students answer to the question with their voice. Only if the system confirms the students' confidence, then the students go one step further. If not, the students have to have another mini-test until the system confirms their confidence. 10 different questions were prepared for mini-test and the final examination with almost same level of SPI tests. Other than these, extra mini-tests were also prepared for extra exercises in case of the system did not confirm the students' confidence or students selected a wrong answer from the candidates.

4.2 Experimental results

Table 1 shows the results from the experiments. The table shows the score and the time required for the learning with 10 mini-test questions and another 10 final examination test questions extracted from SPI test for the existing and the proposed e-learning system with students' confidence evaluation with human voice recognition. It takes 55.4% longer time for the proposed system than the existing system for the mini-tests due to the fact that the proposed system insists the students

another exercises only if the students answer to the question without full confidence. Meanwhile the score of the final examination shows 33.5% improvement on the score. Furthermore, the time required for final examination is reduced by 45.5% so that it may said that the proposed e-learning system makes an effort to improve the score effectively.

5. Concluding remarks

It is found that the proposed e-learning system with students' confidence evaluation improves score for the final examination remarkably; the required for the exercises is much longer compared to the traditional e-learning system without confidence evaluation.

References

(1) International Journal on E-Learning
http://www.aace.org/pubs/IJEL/default.htm
(2) National Technological University
http://www.ntu.edu/
(3) European Community Action Scheme for the Mobility of University Students
http://europa.eu.int/comm/education/index_en.html
(4) University Mobility in Asia and the Pacific
http://www.umap.org/
(5) Shareable Content Object Reference Model
http://www.elc.or.jp/cgi-bin/csvmail/download.htm
(6) Information Technology for Learning, Education and Training
http://jtc1sc36.org/
(7) Advanced Distributed Learning Initiative
http://www.adlnet.org/
(8)Kohei Arai and Hiroshi Yoshida, Method for emotions recognitions of the trainees in e-learning systems, Proceedings of the 29th general conference of the Japanese Society for Information and Systems in Education at Kagawa University, Aug.20-22, 2004.
(9)Lyn Clark, Voice recognition with software applications, Glenco/ Mcgraw Hill Post Secondary, p.508, 2000.
(10)Richard L. Klevans, Voice recognition, Artech House Telecommunications Library, p.171, 1997.