Ontological Descriptions for Eye Movement Data and Mental States in Taking Computer-based Multiple-Choice Tests

Keiichi MURAMATSU[a*], Kazuaki KOJIMA[a] & Tatsunori MATSUI[a]
[a]*Faculty of Human Sciences, Waseda University, Japan*
*kei-mura@ruri.waseda.jp

Abstract: Recently, the research in intelligent educational systems has much interest in exploring data from educational settings to understand learners behavior and mental states. To further understand learners based on such the data and sophisticate supportive interventions by educational systems, knowledge of relationships between learners' behavior and mental states must be shared. To propose a framework to share the knowledge, this study attempted ontological descriptions for learners' eye-movement and mental states in solving computer-based multiple-choice problems. The current study forms a technical basis for development of an IMS (Intelligent Mentoring System) in which an automatic mentoring function is implemented with ITS (Intelligent Tutoring System).

Keywords: Ontology, mental state, eye movement, computer-based test

1. Introduction

Recently, data such as operation logs, face images, eye movements and various physiological indices has been analyzed to recognize situation and state of learner who uses e-learning system. Broad-ranging analyses of the data can provide understandings of learners' mental states in addition to their knowledge structure. For example, affective states such as confidence or confusion [1, 7], unusual states such as impasses or illusion [2, 13] and subjective difficulties for problems [9] are tried to be recognized.

In these studies, information of eye movement is one important data generally used and is useful for realizing human mental processes. Due to higher sampling frequency than time interval of operation [9] and required time for learning [13], such data from eye movement enables to detect learners' mental states in more detail [6, 12]. On the contrary the disadvantage of focusing on that information is that device for the measurement of eye movement is not available in ordinary learning environment. However, knowledge of the eye movement is helpful in understanding of learners' mental states, when it is considered that fixation position can be estimated from face image obtained through stereo-camera [10] which has the potential to become a common device.

Knowledge of learners' behavior such as eye movement is helpful for building intelligent system which supports learning from both aspects of knowledge and mental states. We named such the system as Intelligent Mentoring System (IMS). One of its main characteristics is diagnostic function of learner model considering mental states of learners. Because mental states can instantly change in a short activity (e.g., solving of a single problem), IMS is required to monitor learners at all time and give feedback based on diagnosis. The IMS provides integrative learning-support including real-time estimation of learners' mental states and selection of ways to support learners, in addition to diagnosis of learners' knowledge structures and determination of teaching strategies provided by ITS (Intelligent Tutoring System). Figure 1 shows skeleton framework of IMS.

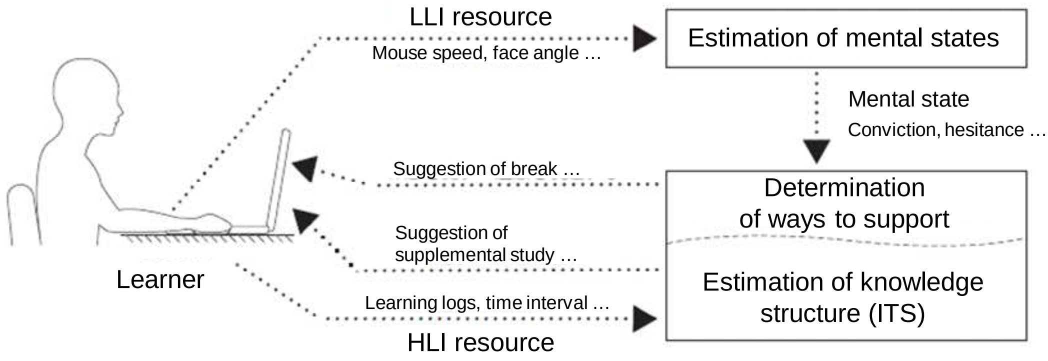


Figure 1. Skeleton framework of Intelligent Mentoring System

Interactions between user and system can be captured in levels of cognitive activity; high-level interactions (HLI) and low-level interactions (LLI). The HLI is explicitly accompanied with user's awareness and is consequently illustrated by data resource which is sampled in large grain sizes. On the other hand, LLI is not always accompanied with user's awareness and is consequently illustrated by data resource which is sampled in very smaller grain sizes. For example, learning logs, time interval of operation [9], and required time for learning [13] can be the former data resources, and moving speed of a mouse, face or posture angle of learner, and gaze position [10] or eye movement of learner can be the latter data resources. Focusing on LLI resources, IMS aims at estimation of learners' mental states from their unarticulated and semi-conscious behaviors.

In e-learning environment, operations of learners who take a multiple-choice test with mouse click are regarded as HLI resource, and behaviors of learners who gaze problem and selection statements are regarded as LLI resources. IMS function to determine ways to support learners is required to integrate information about knowledge structure estimated from HLI resources (this corresponds to function of ITS) and mental states estimated from LLI resources.

In the current study, our purpose is to develop framework to integrate the information and provide consistent description of knowledge implemented in IMS. We focus LLI resource and mental states of learners, because the knowledge about relationships among them is task-oriented and independent from knowledge structures of specific study domains. For this purpose, we build an ontology for description of relationship between eye movement of learner and associated mental states during taking multiple-choice tests as an example of the knowledge. First, we introduce experiment to obtain eye movement data. Second, we show ontological descriptions of multiple-choice test event in which eye movement of learner and associated mental states are positioned.

2. Gaze on Multiple-Choice Tests

2.1 Experiment to Obtain Eye Movement Data

We conducted an experiment to obtain eye movement data of learners during taking multiple-choice test and subjective measurement data about their mental states. Mental states of the learners are extracted from questionnaires. In this experiment, participants were asked to answer thirty four-choice tests which require only encyclopedic knowledge. To observe learners' behavior to read texts and retrieve answers, we created these tests which do not require learners of mental processing and reasoning.

In this experiment, the tests and questionnaires were displayed on a PC monitor full-screen by a program implemented by the second author. First, a button labeled "*proceed next*" was displayed in a position where a problem statement appeared. After clicking of a mouse button on the displayed button, a four-choice test was appeared (Figure 2). When the mouse cursor entered in areas of selection statement, the statement turned red. The red statement was selected as the response with clicking the mouse button on it. Second, questionnaires about the test were displayed after selection of the response. The first questionnaire asked whether the participants respond the selection that they had judged to be the answer. The second asked how the answer was familiar by selecting one of "I know it, I could answer without the choices (*recall*)", "I remembered the answer from the choice (*recognition*)", "I don't know, but guessed from the choices (*guesstimate*)", or "I had no idea (*no-idea*)". The third asked how the each selection was close to the answer. Every selection was evaluated with "it's definitely the answer/ not the answer", "it's probably the answer/ not the answer", or "I cannot judge".

Five undergraduate students participated in this experiment. The participant was seated in front of a desk where a PC monitor (the resolution was 1280 * 1024 pixels) was set up and was asked to answer four-choice tests with a mouse as quickly and correctly as possible. After answering of each test, s/he responded to the questionnaires. As practice task, s/he answered one test and its questionnaires prior to answering the thirty problems. During the task, each participant's eye movements were recorded with EMR-AT VOXER manufactured by NAC Image Technology Inc. Eye movement data included values of x and y coordinates on the screen (pixels) which are sampled at 60 frames/sec.

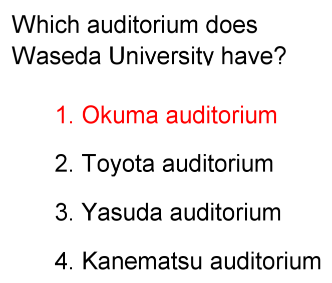


Figure 2. Example of four-choice test displayed on a PC monitor

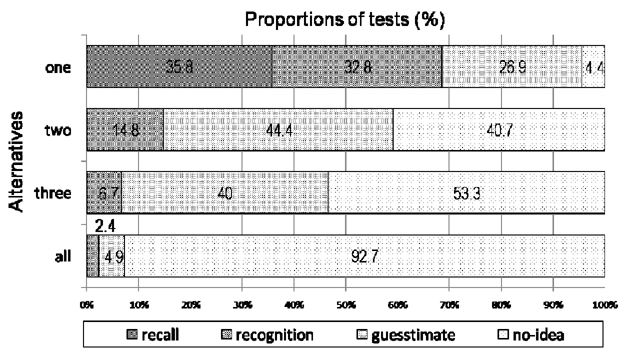


Figure 3. Proportions of tests in each category of alternatives

2.2 Categorization of Responses to Tests

According to the number of response definitely not to be the answer in the third questionnaire, possible selections in each test was decided as alternatives. For example, one alternative remains when three selections were evaluated not to be the answer, and all alternatives remain when nothing was evaluated not to be the answer. Categories of response to tests were decided by combining the alternatives and familiarities of the second questionnaire. Figure 3 shows the proportions of these categories.

However, eye movement data of two participants was removed in analysis because of deficits by mechanical problems. After error processing and smoothing, eye movement data for 58 tests used in analysis. Then the categories were made as follows; 14 tests were one/recall, 12 tests were one/recognition, 6 tests were one/guesstimate, 6 tests were two/guesstimate, 4 tests were two/no-idea and 16 tests were all/no-idea.

2.3 Gaze Patterns

We analyzed gaze patterns visualizing eye movement data. The problem statement was positioned at 96-256 pixels in y-axis direction and selection statement from 1 to 4 were positioned 352-440 pixels, 496-584 pixels, 640-728 pixels and 784-872 pixels respectively. Therefore, plotting values of y coordinate against frames indicates participants' gaze pattern during taking a test.

We found three typical gaze patterns. Figure 4 (a) shows an example of a pattern looking-ahead at selection statement. This pattern was observed in 15 tests and indicates that the participant looked-ahead at selection during reading problem statement. Figure 4 (b) shows an example of a pattern looking-back at problem statement. This pattern was observed in 16 tests and indicates that the participant looked-back at problem statement while s/he scans all selection statements. Figure 4 (c) shows an example of a pattern looking-back at selection statement. This pattern was observed in 11 tests and indicates that the participant looked-back at the previous selection statement while s/he scans all selection statements.

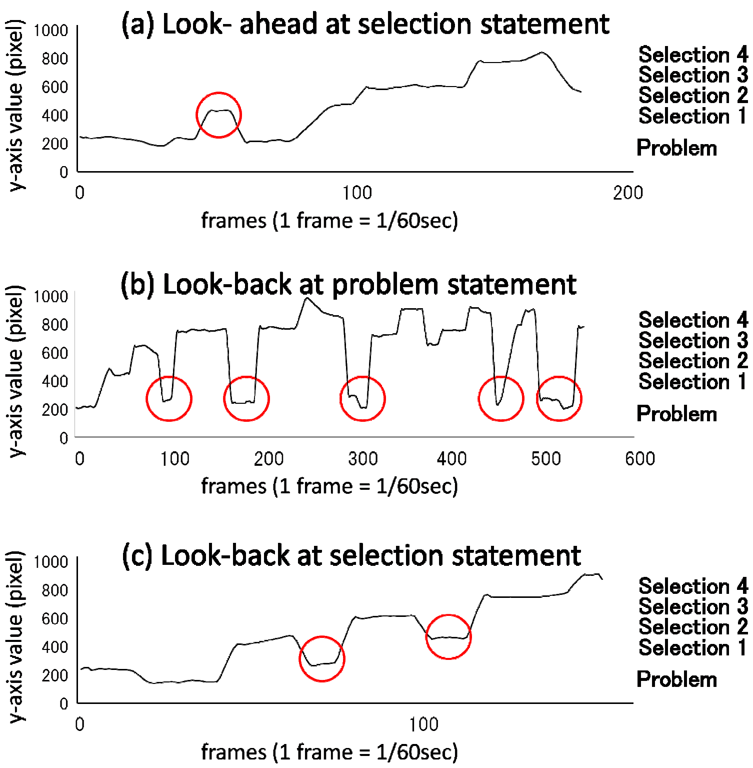


Figure 4. Typical gaze patterns

Table 1. Proportions of gaze patterns and categories of response

	look-ahead at selection	look-back at problem	look-back at selection	cases
one/recall	21% (3)	29% (4)	21% (3)	14
one/recognition	25% (3)	42% (5)	25% (3)	12
one/guesstimate	17% (1)	50% (3)	17% (1)	6
two/guesstimate	34% (2)	1 7%(1)	34% (2)	6
two/no-idea	0	0	0	4
all/no-idea	38% (6)	19% (3)	13% (2)	16

2.3 Relationships between Eye Movement and Mental States

Proportions of gaze patterns and categories of response are summarized as Table 1. Proportion and number of cases which are observed to be each gaze pattern are shown by row. The categories of familiarity based on the second questionnaire indicate learners' mental state called conviction, and the number of alternative is interpreted to lead learners' hesitance in selection.

According to this summary, the pattern of "look-ahead at selection" marked high value of proportion at two/guesstimate and all/no-idea. Thus, this pattern seems to be closely-linked to low conviction and high hesitance. The pattern of "look-back at problem" marked high value of proportion at one/recognition and one/guesstimate, so that this pattern seems to be closely-linked to middle conviction and low hesitance. Finally, the pattern of "look-back at selection" marked high value of proportion at only two/guesstimate, and so this pattern seems to be closely-linked to middle conviction and hesitance.

3. Ontological Descriptions

3.1 Ontological Engineering

Ontological engineering is one of methodologies to describe knowledge systematically. From knowledge-based viewpoint, "ontology is defined as a theory (system) of concepts/vocabulary used as building blocks of an information processing system" [4]. In ontology development environment Hozo[1], each node represents a whole-concept and has some slots which represent part-of or attribute-of relations. Hozo supports describing role concepts which represents a role depends on contents of each whole-concept. For example, a teacher role which is played by a human only in a context of school, and he does not play the role out of the school. In other words, every slot has a role under whole-concept implying a context. In the context, a class of instances which can play a role is defined by a class constraint, and it is called role-holder [3].

We referred a top-level ontology YAMATO[2] [6]. According to YAMATO, an entity is divided into three classes of a physical, abstract and semi-abstract. While instances of physical class need 3D space and time to exist, instances of abstract class needs neither of them. Instances of semi-abstract class need only time to exist, and the class contains mind, representation, a content and a representation form. Representation such as novels, poems, paintings, music, and symbols is distinguished from its proposition and form of representation [5]. Representation has a content slot which indicates a role played by a proposition, and has a form slot which indicates a role played by a representation form.

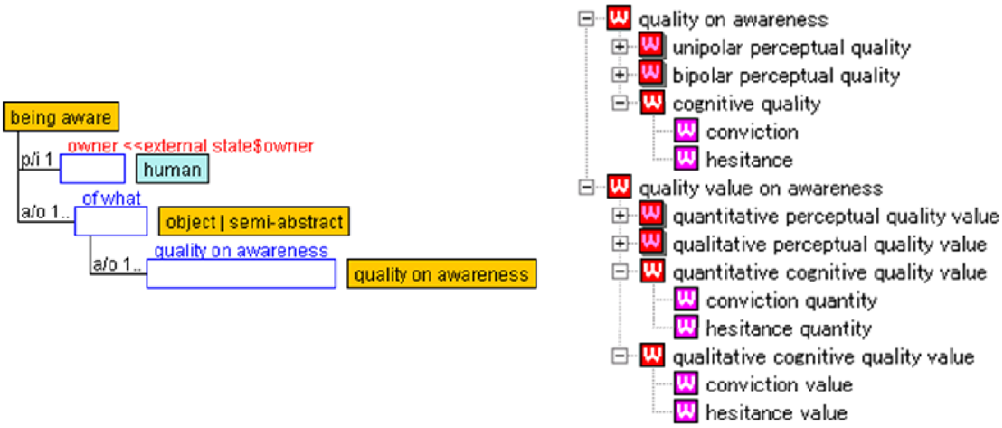
3.2 Attributes on awareness

We have partially expanded YAMATO to describe a subjective evaluation which is regarded as an expression of a psychological quantity. More precisely, it is defined as a representation of quality (defined in YAMATO) based on doer's awareness [8]. Doer's awareness is described in a state "being aware" (figure 5a). This is defined subclass of "external state" in YAMATO. Objects of awareness are represented by "of-what" role-holder played by a physical or a semi-abstract. Subslot of of-what is "cognitive quality"

[1] http://www.hozo.jp

[2] http://www.ei.sanken.osaka-u.ac.jp/hozo/ontolibrary/upperOnto.htm

played by “quality on awareness” represents a psychological quality which a doer subjectively feels.



(a) State of “being aware” (b) Hierarchy of quality and quality value on awareness

Figure 5. State of “being aware” and quality on awareness

Qualities on awareness and their values are sharply distinguished from physical qualities and their values defined in YAMATO. Figure 5b shows hierarchy of “quality on awareness” and “quality value on awareness”. Learners’ psychological quality such as conviction and hesitance are defined as subclass of “cognitive quality” under the quality on awareness. For example, “conviction” has two “referring to” slots; one is played by “conviction quantity” and the other is played by “conviction value”. The conviction quality is subclass of “quantitative cognitive quality value”, and the conviction value is subclass of “qualitative cognitive quality value”. Both quantitative and qualitative cognitive quality values are defined under “quality value on awareness”.

3.3 Multiple-choice test event

3.3.1 Description of learner’s eye movement

To clarify relationships between learners’ eye movement and mental states, we defined “multiple-choice test event” which represents that a learner takes multiple-choice test in an e-learning environment (figure 6). This is subclass of “extrinsic accomplishment” which is defined under “ordinary event” in YAMATO, and consists mainly of participants (objects) and constituting processes (actions). A learner and a learning material which participate in this event are defined as a role of participant inherited from “occurrent” defined in upper level. The learner role is played by a “human”, and learning material is played by a “representing thing”. The representing thing is a physical thing which mediates "representation" and is defined to have a “representation media” which is played by a “physical” thing in YAMATO. Thus, the learning material also consists of representations of a test and a representation media which is characterized by positions of a text and choices on the multiple-choice test.

A “constituted by” slot of multiple-choice test event, which is inherited from ordinary event, is played by “solve” action. In YAMATO, physical action is defined to be composed of “bodily motion”. Thus, we defined “gaze” as a bodily motion and composed the solve action with gaze motions in the context of multiple-choice test. These slots represent “look-ahead at choice”, “look-back at text” and “look-back at choice” as role-holders which are played by gaze motions. These slots also defined to have “fixation position” played by positions of a text and choices on the multiple-choice test.

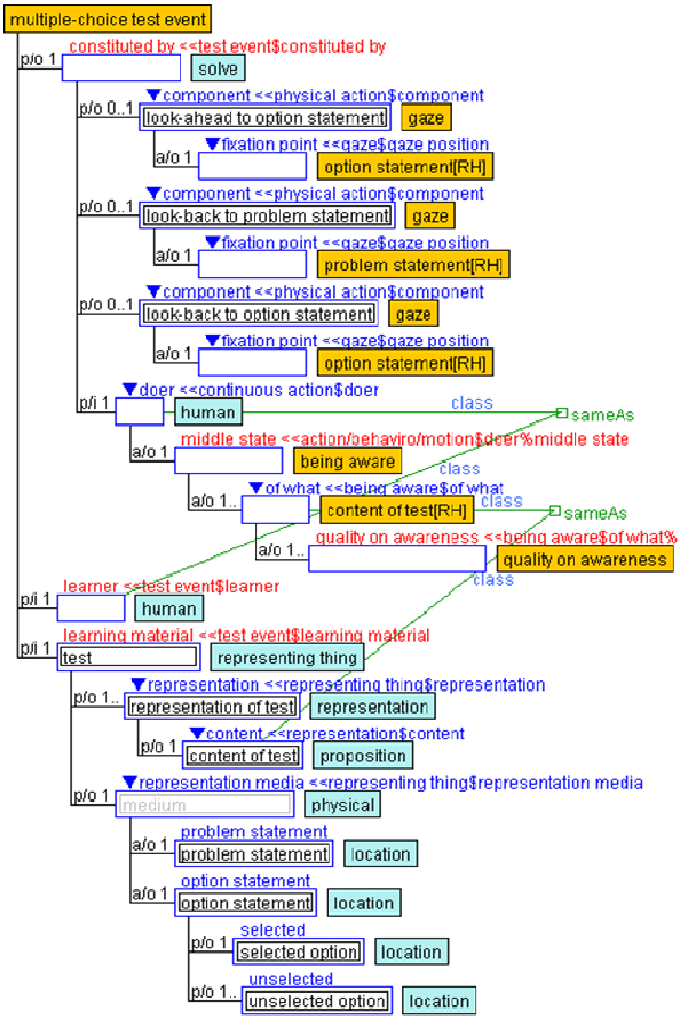


Figure 6. Description of multiple-choice test event

3.3.2 Description of learner's mental states

In the "constituted by" slot which is mentioned above, "doer" played by human represents a subject of the action to solve and motion to gaze. The doer's "middle state" is played by "being aware" and its "of what" slot represents content of consciousness. In this context, the content of consciousness is "content of test" which is a role-holder played by a proposition and is linked to subsubslot of the learning material with a "same as". Learners' mental states are described as "quality on awareness" role played by a "quality on awareness" which represents a psychological quality.

As mentioned above, conviction and hesitance are defined in the current study and can be player of the quality on awareness role. Thus, relationships between learner's eye movement and mental states are clarified in the multiple-choice test event.

4. Discussion and Conclusion

In the current study, we developed an ontology to provide consistent descriptions of knowledge for IMS. We described multiple-choice test event and positioned learner's eye movement and mental states in it. These descriptions can support to identify knowledge about relationships between eye movement and mental states which were obtained through empirical data. In a practical manner, when the gaze behaviors we found in section 2 are observed, each behavior is represented by setting cardinality of the role-holders played by gaze (figure 6) to 1. Along with this, learner’s mental states are represented by playing the quality on awareness role by conviction or hesitance. Hence, the ontological descriptions proposed in the current study form the outline of framework to manage knowledge implemented in IMS. In future work, we expand description of mental states and LLI resources such as easy/difficult, boring/interesting, confused/comprehending and tired/concentrating estimated facial image and face angle [11]. The main issues that must be addressed are how to clarify semantic constraints among various mental states and LLI resources.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 22300294.

References

[1] Arroyo, I., Cooper, D. G., Burleson, W., Woolf, B. P., Muldner, K. & Christopherson, R. (2009). Emotion Sensors Go To School. *Proceeding of the 14th International Conference on Artificial Intelligence in Education* (pp. 17-24). IOS Press.

[2] Horiguchi, Y., Kojima, K., & Matsui, T. (2009). A Study for Exploration of Relationships between Behaviors and Mental States of Learners For an Automatic Estimation System. *Proceedings of the 17th International Conference on Computers in Education* (pp. 173-176). Hong Kong: Asia-Pacific Society for Computers in Education.

[3] Kozaki, K., Kitamura, Y., Ikeda, M., & Mizoguchi, R. (2000). Development of an environment for building ontologies which is based on a fundamental consideration of ”relationship” and ”role”. *Proceedings of the Sixth Pacific Knowledge Acquisition Workshop* (pp. 205-221). Sydney: University of New South Wales.

[4] Mizoguchi, R., Vanwelkenhuysen. J., & Ikeda. M. (1995). Task ontology for reuse of problem solving knowledge. *Proceedings of Knowledge Building & Knowledge Sharing 1995* (pp. 46-59) .

[5] Mizoguchi, R. (2004). Tutorial on ontological engineering - Part 3: Advanced Course of Ontological Engineering. *New Generation Computing*. 22(2), pp. 193-220.

[6] Mizoguchi, R. (2010). YAMATO: Yet Another More Advanced Top-level Ontology, *Proceedings of the Sixth Australasian Ontology Workshop* (pp. 1-16). Sydney : ACS.

[7] Muldner, K., Christopherson, R., Atkinson, R. & Burleson W. (2009). Investigating the Utility of Eye-Tracking Information on Affect and Reasoning for User Modeling. *Proceedings of the 17th International Conference on User Modeling, Adaptation and Personalization* (pp. 138-149). Berlin: Springer.

[8] Muramatsu, K., Togawa, T., Kojima, K,. & Matsui, T. (2011). Proposal of A Framework to Share Knowledge About Consumer’s Impressions. *Proceedings of the 3rd International Conference on Agents and Artificial Intelligence*. Vol. 1 (pp.388-393) SciTePress.

[9] Nakamura, Y., Akamatsu, N., Kuwabara, T., & Tamaki, M. (2002). Method for Detecting a Learner's Stalled Situation Using a Dispersion of Operation Time Intervals. *IEICE Transactions on Information and Systems*, J85-D (1), 79-90 (in Japanese).

[10] Nakamura, K., Kakusho, K., Murakami, M., & Minoh, M. (2010). Estimating Learners' Subjective Impressions of the Difficulty of Course Materials by Observing Their Faces in e-Learning. *IEICE Transactions on Information and Systems*, J93-D (5), 568-578 (in Japanese).

[11] Nosu, K. & Kurokawa, T. (2006) A Multi-Modal Emotion-Diagnosis System to Support e-Learning, *Proceedings of the First International Conferenceon Innovative Computing, Information and Control* Vol. 2 (pp. 274-278). IEEE Computer Society.

[12] Salvucci, D. D. (2001). An Integrated Model of Eye Movements and Visual Encoding, Journal of Cognitive Systems Research, 1(4), 201-220.

[13] Ueno, M. & Nagaoka, K. (2005). On-Line Analysis of e-Learning Time based on Gamma Distributions. *Proceeding of World Conference on Educational Multimedia, Hypermedia & Telecommunications* (pp. 3629-3637). AACE.