Development and Research of an Affective Learning System Combined with Motion-Sensing Interaction, Augmented Reality, and Mid-Air Projection

Hao-Chiang Koong Lin[1], Wen-Chun Hsu[1,2], Tao-Hua Wang[3], Yu-Chun Ma[1], Meng-Chun Tsai[1]

[1] Department of Information and Learning Technology, National University of Tainan, Taiwan
[2] Department of Visual Communication Design, Kun Shan University, Taiwan
[3] Department of Engineering Science, National Cheng Kung University, Taiwan
{koonglin, kenny.hsu0118, wangsab, coldlavender}@gmail.com, cherkimo@yahoo.com.tw

Abstract

The present study developed a novel learning system by using a combination of motion-sensing interaction, augmented reality, and mid-air projection technologies. To measure the system usability, user satisfaction of system interactions, and the relationship between users' operation behaviors and their emotions for users of different learning styles, the present study used the VAK learning styles questionnaire, the system usability scale (SUS), the questionnaire for user interaction satisfaction (QUIS), and sequence analysis. A total of 43 users from Tainan, Taiwan participated in this study.

The experiment results showed that the SUS had a mean score of 73.8 (indicating favorable system usability) and that the user satisfaction scale had a mean score of 5.46 (which was higher than the mean of 4 on the 7-point scale, indicating that the users were subjectively satisfied with the system's human-computer interactions). The present study obtained the following results: On the whole, the users were more likely to change from courses to games when they experienced negative emotions; users with a visual learning style were likely to switch between courses and games when they experienced negative emotions; the vast majority of the operation behaviors of users with an auditory learning style were categorized as "no emotions" because the teaching materials that were favorable to such a learning style were insufficient; and that the operation behaviors of users with a kinesthetic learning style were relatively more diversified, in which none of the operation behaviors accounted for a high percentage.

Keywords: Motion-sensing interaction technology, Augmented reality, Mid-air projection, Affective learning system

1 Introduction

Because of the constant changes in digital technology and the diversification of teaching media and learning aids, learning contents and the presentation of teaching materials have become increasingly rich. Recent studies have shown that by combining motion-sensing technology with augmented reality technology in teaching [1-2], users' learning motivation and participation can be elevated and superior learning results can be achieved. Concerning positive and negative emotions, the former facilitates a smooth learning process [3], whereas the latter leads to learning difficulties. To enable computers to recognize users' emotions, the present study introduced an emotion recognition mechanism to a system and used learning interactions to improve users' emotions and their learning results [4].

The present study developed a novel learning environment by combining motion-sensing technology with augmented reality and mid-air projection technologies. In addition, the present study presented an affective learning system, in which an emotion recognition mechanism was employed to facilitate user interactions. The goals were to "manage" user emotions to elevate their involvement, learning motivation, and learning participation as well as to identify the behaviors of users with different learning styles. Accordingly, this study explored the following questions:

(1) What was the usability of the affective learning system combined with motion-sensing, augmented reality, and mid-air projection technologies?

(2) How satisfied were the users with the user interactions provided by the system?

(3) What was the relationship between users' operation behaviors and their emotions?

(4) What were the differences between users of different learning styles in terms the relationship between their operation behaviors and their emotions?

*Corresponding Author: Wen-Chun Hsu; E-mail: kenny.hsu0118@gmail.com
DOI: 10.3966/160792642018111906031

2 Literature Review

2.1 Affective Computing

Affective computing analyzes changes in people's physiological signals (e.g., words, voices, electroencephalography [5] and facial expressions) to understand people's emotions and make appropriate responses [6]. Of the many methods that can be used to recognize emotions, facial expression recognition is commonly used. The driver fatigue can be detect by eyes pattern or motion ian image processing [7], and multiple facial expression recognition methods have already been developed [8-10] and even to understand a combination of different emotions [11].

Emotion is one of the key factors influencing people's learning. Positive emotions enhance people's problem-solving abilities [3]. Many studies have already attempted to elevate users' learning results by influencing their emotions, and affective tutoring systems are considered a major tool that can influence such emotions [12-15]. Affective tutoring systems detect users' learning status and emotional state to provide appropriate and timely feedback [16]. Ammar, et al. [17] detected and determined users' emotions by using their facial expressions. Alepis and Virvou [18] combined ontology with emotions expressed by system users to explore the teaching strategies to be used for different user emotions. Graesser, et al. [19] built a tutoring system with affective factors that successfully elevated users' learning results. Sarrafzadeh, et al. [20] provided a set of emotion-oriented counselling for elementary school math classes. Wang, et al. [21] introduced a system that collected users' emotions that they actively express in the learning process to determine what type of feedback should be given by the system design agents. Calvo and D'Mello [22] conducted a survey described recent developments in the field of affective computing with emphasis on affect detection, and it explicitly explored the multidisciplinary foundation and provided meta-analyses on current reviews of affect detection systems that focus on traditional affect detection modalities like physiology, face, and voice, and also discussed emerging research. Therefore, this research referred to the survey to develop facial expression recognition module in order to detect emotion with the emotion classifiers. Misra and Saha [23] studied differs from the related existing works, which focus on detecting emotions of the users based on their activities.

2.2 Motion-Sensing Interaction Technologies

Because of the advances in technology, human-computer communication has evolved from being human-machine interface-based (which involves the use of a keyboard and a mouse) to being natural user interface-based (NUI-based) (which involves the use of people's natural body movements). The commercialization of motion-sensing devices such as Microsoft Kinect, Asus Xtion, and Leap Motion has further driven the development and application of motion-sensing interaction technologies. In addition, because these devices generally feature corresponding development tools and convenient development methods, many scholars became involved in studies that utilized motion-sensing interaction technologies. Various natural body movement-based operating methods of motion-sensing technologies have been proposed, such as the body movement method and the hand gesture method [2], and a study presented recently a system for pedestrian tracking and activity recognition in indoor environments using exclusively common off-the-shelf sensors embedded in smartphones[24].

Homer, et al. [2] developed a set of motion-sensing e-book reading systems by using various body movements, in which they found that motion-sensing-based word-learning games significantly improved young children's sight word recognition and that children enjoyed the games very much.

Berri, et al. [25] employed Kinect Sensor to enable robots to locate people's faces and locations and avoid obstacles; the said sensor could also be operated through hand gestures.

2.3 Mid-Air Projections

Mid-air projections use light refractions and reflections to gather light to form three dimensional images. Such projections adopt the basic concept of Pepper's ghost, which uses glasses as mirrors by placing in front of a dark background. Glasses placed at a 45° angle will allow viewers to see images that are perpendicular (i.e., at a 90° angle) to their line of sight. In addition, because glasses are transparent, they allow images from different locations to overlap. Such a principle is widely used in stage performances and in magic shows [26]. Concerning the display of mid-air projections, two main methods are currently used: the first involves projecting the images onto a projection screen made of transparent materials, and the second involves displaying images in a showcase.

2.4 Learning Styles

In this study, the users learned by operating the motion-sensing system using hand gestures. The teaching materials adopted were information presented on the screen using display technologies. Because of the teaching materials and system employed in this study, the visual-auditory-kinesthetic (VAK) learning styles model was used in the present study to categorize users by learning styles. The VAK learning styles model [27] divides users' learning behaviors according to the senses that they used in learning. After filling the VAK learning styles questionnaire, users were divided into users with a visual learning style, users with an auditory learning style, and users with a kinesthetic learning style. Users with a visual learning style like to learn by making observations. This type of users typically achieves learning results by using visually stimulating learning methods. Users with an auditory learning style like to acquire new information through the auditory sense; when they read, they like to read the words out loud or quietly to help them memorize the learning content. They acquire knowledge and analyze data through hearing and achieve learning results by listening to videos and teaching materials presented in classes. Users with a kinesthetic learning style prefer to learn by using physical experience-based learning methods such as participating in teaching activities or undergoing total physical responses. They like to learn teaching materials through physical activities and achieve learning results through interactions and object operations.

From the above literature, it is known that emotion is one of the important factors that affect people's learning, and there has been much research about the technology of emotion detector was applied to various teaching categories, but the most of those researches provide a common learning environment with traditional interface. The biggest difference between this study and former studies is that this study employed the latest motion-sensor technology to provide more intuitive operation and applied Augmented Reality technology and Mid-Air Projections to create a more immersive environment in order to analysis and enhance effectiveness of learners with different learning styles.

3 Research Methods

3.1 Environment Featuring a Combination of Motion-Sensing, Augmented Reality, and Mid-Air Projection Technologies

The present study used the Kinect motion-sensing device and referred to related data and human interface guidelines developed by Microsoft Kinect to select appropriate system operation movements. By using the Microsoft Kinect SDK, the Developer Toolkit, and a mid-air projection device, this study projected the affective learning system onto real world environments. Moreover, this learning system was combined with a motion-sensing device to allow users control system buttons projected onto mid-air with their hand gestures. In addition, anthropomorphic interactive agents were designed to interact with users. Figure 1 shows an environment featuring the combination of motion-sensing, augmented reality, and mid-air projection technologies, in which A is a 24-inch computer screen, B is a transparent acrylic board, C is the Kinect device, and D is a table. Because the system required users to perform the experiment activities while seated, the transparent acrylic board used for mid-air projections was placed at a 45° angle to the computer screen (i.e., the projection device) and within the users' line of sight while they were seated. Also, because images projected by computer screens are reversed left and right as a result of optical mirroring (similar to images seen in mirrors), the present study reversed the images first using the software UltraMon to allow the projected images to be identical to normal system operation images.

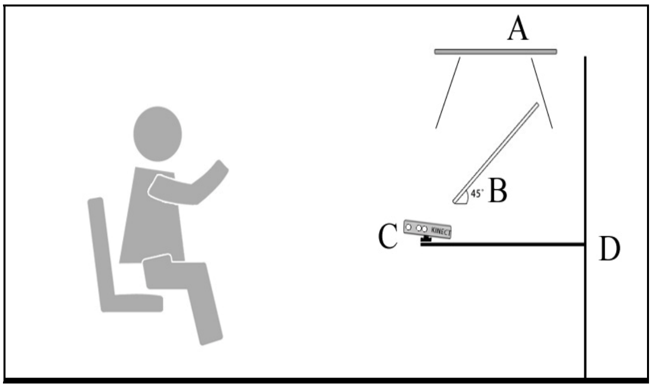


Figure 1. An environment featuring a combination of motion-sensing, augmented reality, and mid-air projection technologies (this image was created by the author of this study)

3.2 Affective Learning System

The system introduced in this study featured two major axes, which were user emotion collection and the learning system. Concerning obtaining users' emotions, two methods were employed: for the first method, the system directly asked the users to indicate their current emotional states; for the second method, the system used the facial expression recognition function. Regarding the learning system, it covered courses, interactive agents, mini games, and system records. The overall system structure contained five modules: course module, facial expression recognition module, interactive agent module, game module, and system records module. Reading is one of the main ways of learning, but reading is a complex cognitive process that is often difficult to observe [28, 29], therefore this study observed and analyzed the conditions of learners' reading by affect detection. Figure 2 shows the system interface, and Figure 3 shows the affective learning system (combined with motion-sensing, augmented reality, and mid-air projection technologies) in operation.

To develop the facial expression recognition module, the system employed open library EmguCV that encapsulated OpenCV components by using C#. EmguCV featured powerful image processing capabilities and many libraries, simplifying system developments and reducing system development time. The facial expression recognition module adopted the following procedure: first, locate user's face; second, use HaarTraining (which provided target detection for OpenCV and could be used to train desired classifiers by adjusting its features) and the trained emotion classifiers (which classified the emotions of joy, anger, surprise, fear, confusion, and sadness) to compare the results obtained (i.e., users' emotions) with existing emotion classifier data; and third, identify the user's emotions by matching his/her emotions with the emotion classifiers.

The present study used the six trained emotion classifiers to identify users' facial features and expressions. The facial expression recognition module followed the following procedures: first, activate the Webcam; second, locate user's face (the objective of this step was to limit the facial expression recognition area to that of the user's face to lower the facial expression recognition range); and third, identify user's facial expressions by using the six trained emotion classifiers. Emotions that matched the emotion classifiers' data would be identified; and those that failed to match any emotion classifier were concluded as "no emotions." Once the emotions were detected, learning system strategies were deployed. During this period, the system continued to run the facial expression recognition.

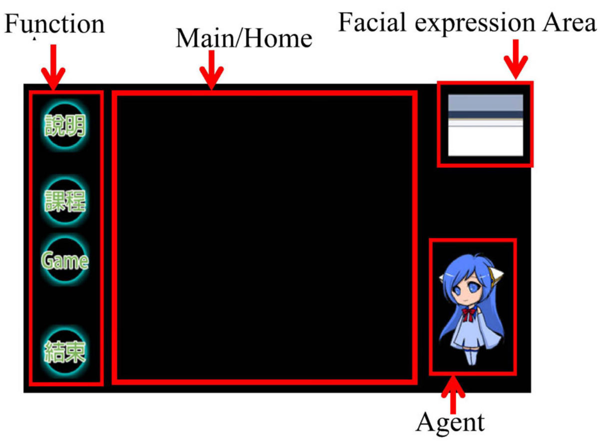


Figure 2. Affective learning system interface

Figure 3. Image of the affective learning system (combined with motion-sensing, augmented reality, and mid-air projection technologies)

3.3 Research Tools

3.3.1 System Usability Scale (SUS)

This study used the system usability scale developed by Digital Equipment Co Ltd. in 1986 to assess users' evaluation of system usability. The system usability scale is reliable, fast, convenient, and low in cost [30]. In this study, the scale contained 10 items/questions and a score was given to each item/question using the five-point Likert scale. The scores ranged from 1 to 5, in which a higher score indicated higher user satisfaction with system usability.

3.3.2 Questionnaire for User Interaction Satisfaction

This study used the questionnaire for user interaction satisfaction (QUIS) developed by the Human-Computer Interaction Lab (HCIL) of the University of Maryland, U.S.A. to evaluate the relationship between human-machine interface and users' subjective satisfaction [31]. The scale measured users' assessments of six system dimensions (i.e., overall reaction of the system, screen display, system terminology and information, learnability, system performance, and usability and user interface) and contained a total of 28 items/questions.

3.3.3 VAK Learning Styles Questionnaire

The present study used the VAK learning styles scale adopted in another study conducted by Wen [32]. The said scale contained 30 items/questions and was used in both studies to assess users' learning styles.

3.3.4 Message Encoding

After the experiment, the users' operation and emotional behaviors (stored in the system records) were matched with the six facial expressions proposed by Ekman and Friesen [33] (i.e., fear, confusion, surprise, sadness, anger, and joy). Next, the emotions were converted into positive or negative emotions by matching them with the two-dimensional emotion model (which divided different emotions into different locations) introduced by Russell [34]. Next, the operation behaviors and emotional states to be observed were coded and sequence analysis was performed, and the flow chart of the whole experiment is shown in Figure 4.

4 Experiment Results

4.1 System Usability Analysis

This study recruited a total of 43 users and analyzed 43 valid samples. The samples yielded a Cronbach's alpha of 0.819, which indicated that the system usability scale demonstrated favorable reliability. The

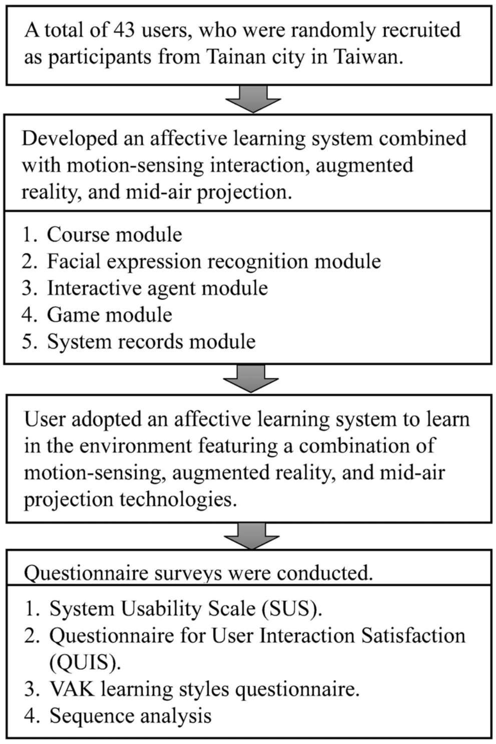


Figure 4. Flow chart of the whole experiment

system usability scale consisted of the following items (Table 1):

(1) Q1. I used this system frequently: the results showed a mean of 3.49 and a standard deviation of 0.83; 48.8% of the users used this system frequently, whereas 7% did not use the system frequently. Approximately 44.2% selected "neither yes nor no." which may be due to the relatively low popularity of motion-sensing mid-air projection environments currently. As a result, the users had a difficult time assigning a score.

(2) Q2. I found this system overly complex: the results showed a mean of 4.09 and a standard deviation of 0.84; 83.8% of the users did not find the system overly complex.

(3) Q3. I found this system very easy to use: the results showed a mean of 4.18 and a standard deviation of 0.85; 81.4% of the users found the system very easy to use.

(4) Q4. I believe that I needed the help from a technical personnel to use this system: the results showed a mean of 3.77 and a standard deviation of 1.09; 67.4% of the users felt that they did not need the help from a technical personnel to use the system, 16.3% of the users indicated that they did, and 16.3% of the users selected "neither yes nor no."

(5) Q5. I found that all the system functions were well integrated: the results showed a mean of 3.77 and a standard deviation of 0.75; 69.7% of the users stated that all the system functions were well integrated, 2.3% of the users felt that the system functions were poorly integrated, and 27.9% of the users selected "neither yes nor no."

(6) Q6. I found many conflicting problems with this system: the results showed a mean of 3.98 and a standard deviation of 0.89; 76.7% of the users indicated that the system worked reasonably well and 18.6% of the users selected "neither yes nor no."

(7) Q7. I think most people would be able to learn to use this system quickly: the results showed a mean of 4.51 and a standard deviation; 93% of the users answered that they could learn to use the system quickly.

(8) Q8. I found this system very difficult to use: the results showed a mean of 4.23 and a standard deviation of 0.84; 83.7% of the users found the system very easy to use.

(9) Q9. I had great confidence in using this system: the results showed a mean of 4.32 and a standard deviation of 0.78; 86% of the users had great confidence in using the system.

(10) Q10. I believe that I need to acquire additional knowledge/skills to be able to use this system: the results showed a mean of 4.06 and a standard deviation of 0.82. 86.1% of the users indicated that they did not need to acquire additional knowledge/skills to be able to use the system.

Table 1. Descriptive statistics obtained using the system usability scale

V	SS	M	SD	% on the 5-point scale (%)					
				1	2	3	4	5	4+5
Q1	43	3.49	.83	2.3	4.7	44.2	39.5	9.3	48.8
Q2	43	4.09	.84	0	7.0	9.3	51.2	32.6	83.8
Q3	43	4.18	.85	0	4.7	14.0	39.5	41.9	81.4
Q4	43	3.77	1.09	2.3	14.0	16.3	39.5	27.9	67.4
Q5	43	3.77	.75	2.3	0	27.9	58.1	11.6	69.7
Q6	43	3.98	.89	2.3	2.3	18.6	48.8	27.9	76.7
Q7	43	4.51	.63	0	0	7.0	34.9	58.1	93
Q8	43	4.23	.84	0	4.7	11.6	39.5	44.2	83.7
Q9	43	4.32	.78	0	2.3	11.6	37.2	48.8	86
Q10	43	4.30	.71	0	0	14.0	41.9	44.2	86.1
Overall	43	4.06	.82	0.92	3.97	17.45	43.01	34.65	77.66

Note. V= Variables, SS=Sample Size, M=Mean, SD= Standard Deviation, and all negatively worded items/ questions had been converted to positively worded items/ questions.

Table 2 shows that concerning the users' assessment of system usability, it had a mean score of 76.63, a median of 77.50, a mode of 67.50, a maximum value of 97.50, a minimum value of 42.50, and a standard deviation of 12.78. With a mean score of 76.63, it indicated that the users thought favorably of the system's usability.

Table 2. Statistical results obtained by converting the scores received on the system usability scale

Sample Size	Mean	Median	Mode	Minimum Value	Maximum Value	Standard Deviation
43	76.63	77.50	67.50	42.50	97.50	12.78

4.2 Analysis of the Users' Interaction Satisfaction

The QUIS showed a Cronbach's alpha of 0.926, indicating favorable scale performance. The QUIS contained six dimensions. Table 3 shows the statistical analysis of six dimensions. All six dimensions displayed a satisfaction score of 5 or above. The overall system mean score was 5.46, and the standard deviation was 0.90. This indicated that the users were subjectively satisfied with the system interactions as well as the human-computer interactions provided by the system.

Table 3. Descriptive statistics obtained using the questionnaire for user interaction satisfaction

	Sample Size	Mean	Minimum Value	Maximum Value	Standard Deviation
Overall reaction of the system	42	5.15	2	7	1.00
Screen display	42	5.66	3	7	.94
System terminology and information	42	5.31	3	7	.86
Learnability	42	5.99	4	7	.79
System performance	42	5.65	4	7	.95
Usability and user interface	42	5.01	4	7	.84
Mean	42	5.46	-	-	.90

4.3 Sequence Analysis

The experiment recruited a total of 43 users, 22, 4, and 13 of whom were users with a visual learning style, users with an auditory learning style, and users with a kinesthetic learning style, respectively. Four of the users did not display a dominant learning style and were thus removed from the analysis.

To understand the relationship between users' operation behaviors and changes in their emotional states, this study coded both the users' operation behaviors and their emotional states (at the time of their operation behaviors) and investigated their operation behaviors (i.e., selected course learning or games/mini-games). Each of the two operation behaviors was divided into three emotional states (i.e., positive emotions, no emotions, and negative emotions) and a total of six codes were assigned. The overall sample then underwent a sequence analysis to explore whether the three different types of users displayed different performances.

4.3.1 Sequence Analysis of the Overall Sample

Table 4 shows the number of users as a percentage of the total number of users for each behavior code. A total of 58% of the users selected courses (39% of whom showed positive emotions), whereas 42% of the users (37% of whom showed positive emotions) selected games.

Figure 5 shows an event change diagram-based sequence analysis of the overall sample. In this diagram, an arrow indicated a significant change and the Z-score showed how significant the event change was; $|Z| \geq 2.58$ indicated a markedly significant change, $|Z| \geq 1.96$ indicated a significant change, and $|Z| < 1.96$ indicated a nonsignificant change [35]. A thicker arrow signified a more significant change. In the sequence analysis, significant changes were only observed in Events A to A, A to B, B to B, C to C, and C to F. On the basis of the emotion change diagram, four results were obtained:

Table 4. The number of users as a percentage of the total number of users for each behavior code

Code	Behavior	The number of users as a percentage of the total number of users for each behavior code (%) (Sample size: 39 users)
A	Selected courses, showing positive emotion	39
B	Selected courses, showing no emotions	13
C	Selected courses, showing a negative emotion	6
D	Selected games, showed a positive emotion	37
E	Selected games, showing no emotions	3
F	Selected courses, showing positive emotion	2

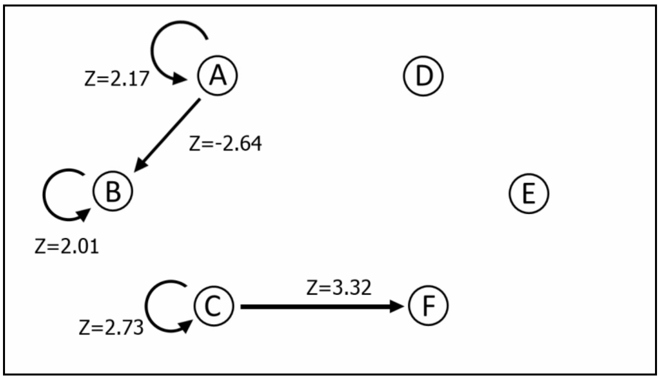


Figure 5. Event change diagram of the overall sample

(1) A considerably high percentage of users displayed positive emotional states during the overall system operation process, as shown in Table 4.6;

(2) At identical emotional states, users were likely to continue learning. In other words, no significant changes in emotions were observed in the users when learning (i.e., no significant changes were observed from Event A (i.e., elected course, showing a positive emotion) to Event B (i.e., selected courses, showing no emotions), or from Event B to Event C (i.e., selected

courses, showing a negative emotion). Events A, B, and C are shown in Figure 5.

(3) Users who experienced negative emotions when they selected courses were likely to switch to games, such as the switch from Event C to Event F (i.e., selected games, showing a negative emotion), as shown in Figure 5.

(4) Users who experienced a change from positive emotions to no emotions (i.e., from Event A to Event B in Figure 5) during the learning process were significantly rare, and $|Z| \geq 2.64$ indicated a significant change; however, the Z value here was negative, indicated that such a change was infrequent.

4.3.2 Sequence Analysis of the Users of the Three Different Learning Styles

Table 5 shows the behavior codes and their respective weights (measured in percentage) for the users of the three different learning styles. For users with a visual learning style, 71% of them selected courses, among which 51% had positive emotions. By contrast, 29% of the users selected games, among which 21% had positive emotions. For users with an auditory learning style, 72% of them selected courses, among which 48% had positive emotions. By contrast, 29% of the users selected games, among which 24% had positive emotions. For users with a kinesthetic learning style, 70% of them selected courses, among which 40% had positive emotions. By contrast, 31% of the users selected games, among which 25% had positive emotions.

Table 5. Weights of the behavior codes (measured in percentage)

Code	Behavior	Behavior codes and their respective weights (measured in %)		
		Users with a visual learning style	Users with an auditory learning style	Users with a kinesthetic learning style
A	Selected courses, showing positive emotion	51	48	40
B	Selected courses, showing no emotions	16	5	21
C	Selected courses, showing a negative emotion	4	19	9
D	Selected games, showed a positive emotion	21	24	25
E	Selected games, showing no emotions	4	5	4
F	Selected courses, showing positive emotion	4	0	2

Figure 6 shows an event change diagram-based sequence analysis of users with a visual learning style. In this diagram, an arrow indicated a significant change and the Z-score showed how significant the event changes were. A thicker arrow signified a more significant change. In the sequence analysis, significant changes were only observed in Events A to B, B to B, B to E, C to C, C to F, and F to C.

The event change diagram of users with a visual learning style was different from that of the overall sample. The differences are listed as follows: first, for users with a visual learning style who showed no emotions, a large proportion switched to games whereas another large portion continued to learn the courses. For example, in Figure 6, both Events B to B and B to E (i.e., selected games, showing no emotions) were observed; two, for users with a visual learning style who showed negative emotions, they were more likely to switch from courses to games. For example, in Figure 6, the switch from Events F to C was significant. Similarly, the switch from Events C to F was significant, showing that when users with a visual learning style experienced negative emotions, a change in such emotions was difficult to achieve; three, the switch from Events A to A became nonsignificant, suggesting that for users with a visual learning style who with a positive emotional state, no clear tendency was identified after they entered course learning; and four, users who experienced a change from positive emotions to no emotions (i.e., from Event A to Event B in Figure 6) during the learning process were significantly rare (i.e., $|Z| \geq 2.06$ indicated a significant change; however, the Z value here was negative, indicated that such a change was infrequent).

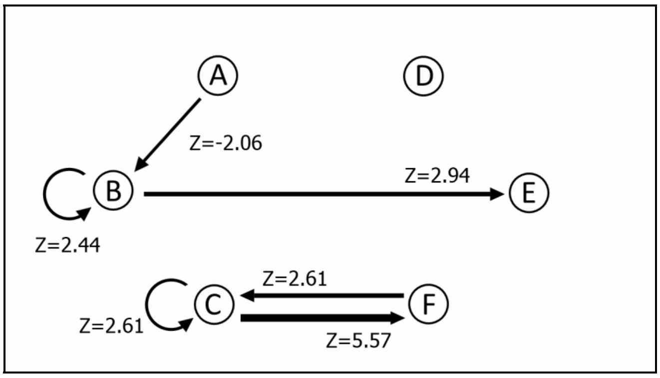


Figure 6. Event change diagram of users with a visual learning style

Figure 7 shows the event change diagram of users with an auditory learning style, in which significant changes were only observed in Events E to B. Concerning the system used in this study, it provided insufficient teaching materials and functions for such a learning style. Both Events E and B entailed "no emotions," indicating that the lack of favorable conditions for the auditory learning style might be the reason that the users engaged in the switch.

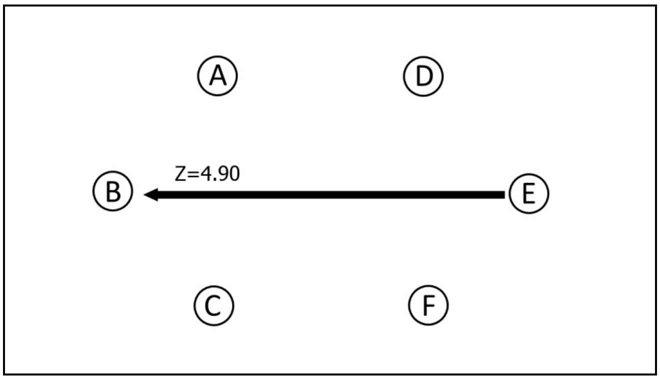


Figure 7. Event change diagram of users with an auditory learning style

Figure 8 shows the event change diagram of users with a kinesthetic learning style, in which no arrows were found. The present study hypothesized that it may be because that all users with a kinesthetic learning style demonstrated unique behavioral trends, resulting in no significant behavioral changes.

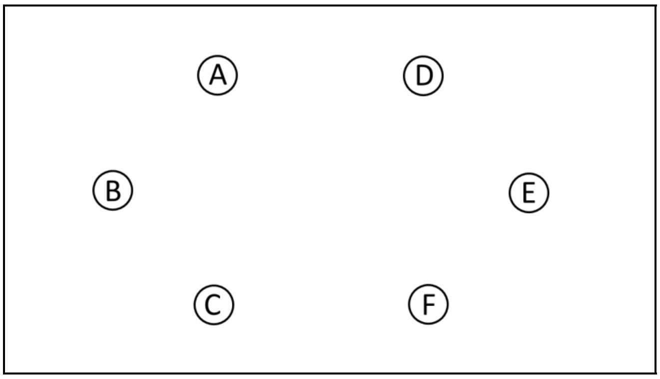


Figure 8. The event change diagram of users with a kinesthetic learning style

5 Conclusion

In this section, research questions that were proposed in this study were answered and the conclusions are as follows:

(1) What was the usability of the affective learning system affective learning system combined with motion-sensing, augmented reality, and mid-air projection technologies?

The analysis showed that the system displayed a usability score of 76.63, which was higher than the scale standard and indicated favorable system usability. Concerning the item "I used this system frequently," a large proportion of the users selected "neither yes nor no." because the result might be due to the relatively low popularity of motion-sensing mid-air projection environments currently. As a result, the users had a difficult time assigning a score.

(2) How satisfied were the users with the user interactions provided by the system?

The analysis showed that the user satisfaction scale had a mean score of 5.46 (which was higher than the mean of 4 on the seven-point scale, indicating that the users were subjectively satisfied with the system's human-computer interactions).

(3) What was the relationship between users' operation behaviors and their emotions?

The present study hypothesized that users who experienced negative learning emotions would switch to games. This hypothesis was verified by the event change diagram-based sequence analysis, which showed that the majority of the users who experienced negative emotional states when learning courses switched to games.

In addition to the aforementioned results, two additional findings were identified from the event change diagram: one, when users continued to engage in the learning process, their emotions mostly remained unchanged, suggesting emotional continuity; and two, users who experienced a change from positive emotions to no emotions during the learning process were rare.

(4) What were the differences between users of different learning styles in terms the relationship between their operation behaviors and their emotions?

The event change diagram of users with a visual learning style showed the following results: those who experienced a change from positive emotions to no emotions during the learning process were rare; those who had no emotions showed two tendencies, the first of which involved continuing their studies, and the second of which involved switching to games; and those who experienced negative emotions continually switched between courses and games, revealing that the course content and games were unable to spark the users' interest.

The event change diagram of users with an auditory learning style showed the following results: the vast majority of the operation behaviors of users with an auditory learning style were categorized as "no emotions." The researcher inferred that such a behavioral model was caused by insufficient favorable learning conditions for users with an auditory learning style.

The event change diagram of users with a kinesthetic learning style showed that none of the users' operation behaviors accounted for a high percentage, which may be because that all users with a kinesthetic learning style demonstrated unique behavioral trends, resulting in no significant behavioral changes.

Concerning future prospects, as there has been some research showed that gender differences had substantial and considerable influence on learning by reading [36], so this research may explore the influence of gender differences in using the affective learning system in the future. Moreover designers may improve visual presentation by designing system interfaces and teaching materials using animations and multimedia. Regarding system operations, because the users in the present study could only drag images by pushing or clenching their fists, more hand gestures or even other motion-sensing devices and speech-to-text recognition system [37-38] may be introduced in order to be applied in wider learning field in the future. With respect to emotion recognition, in addition to facial expression recognition, future researchers may attempt to use additional emotion recognition mechanisms to enhance emotion recognition. Concerning interactive agent interactions, interactive agent games may be added to improve such interactions. Regarding game modules, future researchers may incorporate course content into games to increase the connection between courses and games and make them more valuable. Finally, sound-based education may be added to the learning materials; a user had indicated the wish for a system that could read the course text out loud during classes. Such a recommendation may be realized in the future to produce superior learning results.

Acknowledgements

The authors wish to acknowledge the help of Mr. Chang-You, Chen in experimentation and data analysis, and thank the reviewers and editors for their precious comments.

References

[1] N. Akazawa, Y. Takei, Y. Nakayama, H. Kakuda, M. Suzuki, A Learning Support System for 9× 9 Multiplication table with Kinect, *2013 IEEE 2nd Global Conference on Consumer Electronics*, Tokyo, Japan, 2013, pp. 253-257.

[2] B. D. Homer, C. K. Kinzer, J. L. Plass, S. M. Letourneau, D. Hoffman, M. Bromley, Moved to Learn: The Effects of Interactivity in a Kinect-based Literacy Game for Beginning readers, *Computers & Education*, Vol. 39, No. 4, pp. 378-393, May, 2014.

[3] V. Eyharabide, A. Amandi, Ontology-based User Profile Learning, *Applied Intelligence*, Vol. 36, No. 4, pp. 857-869, June, 2012.

[4] R. M. Felder, L. K. Silverman, Learning and Teaching Styles in Engineering Education, *Engineering Education*, Vol. 78, No. 7, pp. 674-681, April, 1988.

[5] R. Du, R. M. Mehmood, H.-J. Lee, Alpha Activity during Emotional Experience Revealed by ERSP, *Journal of Internet Technology*, Vol. 15, No. 5, pp. 775-782, September, 2014.

[6] R. W. Picard, R. Picard, *Affective Computing*, The MIT Press, 1997.

[7] B. G. Lee, S. J. Jung, W. Y. Chung, Real-time Physiological and Vision Monitoring of Vehicle Driver for Non-intrusive Drowsiness Detection, *IET Communications*, Vol. 5, No. 17, pp. 2461-2469, November, 2011.

[8] T. F. Cootes, C. J. Taylor, D. H. Cooper, J. Graham, Training Models of Shape from Sets of Examples, *British Machine Vision Conference 1992*, London, England, 1992, pp. 9-18.

[9] I. Ari, A. Uyar, L. Akarun, Facial Feature Tracking and Expression Recognition for Sign Language, *2008 23rd International Symposium on Computer and Information Sciences*, Istanbul, Turkey, 2008, pp. 1-6.

[10] P. Metri, J. Ghorpade, A. Butalia, Facial Emotion Recognition Using Context Based Multimodal Approach, *International Journal of Emerging Sciences*, Vol. 2, No. 1, pp. 171-183, March, 2012.

[11] Y.-H. Lee, Virtual Representation of Facial Avatar through Weighted Emotional Recognition, *International Journal of Internet Protocol Technology*, Vol. 10, No. 1, pp. 30-35, March, 2017.

[12] K. H. C. Lin, C. H. Wang, C. J. Chao, M. K. Chien, Employing Textual and Facial Emotion Recognition to Design an Affective Tutoring System, *Turkish Online Journal of Educational Technology*, Vol. 11, No. 4, pp. 418-426, October, 2012.

[13] K. H. C. Lin, C. H. Wu, Y. P. Hsueh, The Influence of Using Affective Tutoring System in Accounting Remedial Instruction on Learning Performance and Usability, *Computers in Human Behavior*, Vol. 41, pp. 514-522, December, 2014.

[14] K. H. C. Lin, C. J. Chao, T. C. Huang, From a Perspective on Foreign Language Learning Anxiety to Develop an Affective Tutoring System, *Educational Technology Research & Development*, Vol. 63, No. 5, pp. 727-747, October, 2015.

[15] K. H. C. Lin, S. H. Su, C. C. Ju, H. C. Yen, T. S. Chin, Construction of Multi-mode Affective Learning System: Taking Affective Design as an Example, *Journal of Educational Technology & Society*, Vol. 19, No. 2, pp. 132-147, April, 2016.

[16] X. Mao, Z. Li, Agent Based Affective Tutoring Systems: A Pilot Study, *Computers & Education*, Vol. 55, No. 1, pp. 202-208, August, 2010.

[17] M. B. Ammar, M. Neji, A. M. Alimi, G. Gouardères, The Affective Tutoring System, *Expert Systems with Applications*, Vol. 37, No. 4, pp. 3013-3023, April, 2010.

[18] E. Alepis, M. Virvou, Automatic Generation of Emotions in Tutoring Agents for Affective e-learning in Medical Education, *Expert Systems with Applications*, Vol. 38, No. 8, pp. 9840-9847, August, 2011.

[19] A. C. Graesser, S. Lu, G. T. Jackson, H. H. Mitchell, M. Ventura, A. Olney, M. M. Louwerse, AutoTutor: A Tutor with Dialogue in Natural Language, *Behavior Research Methods*, Vol. 36, No. 2, pp. 180-192, May, 2004.

[20] A. Sarrafzadeh, S. Alexander, F. Dadgostar, C. Fan, A. Bigdeli, How Do You Know that I Don't Understand? A Look at the Future of Intelligent Tutoring Systems, *Computers in Human Behavior*, Vol. 24, No. 4, pp. 1342-1363, July, 2008.

[21] C. Y. Wang, G. D. Chen, C. C. Liu, B. J. Liu, Design an Empathic Virtual Human to Encourage and Persuade Learners in e-learning Systems, *The first ACM International Workshop on Multimedia Technologies for Distance Learning*, Beijing, China, 2009, pp. 27-32.

[22] R. A. Calvo, S. D'Mello, Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications, *IEEE Transactions on Affective Computing*, Vol. 1, No. 1, pp. 18-37, January, 2010.

[23] S. Misra, B. K. Saha, On Emotional Aspects in Mission-Oriented Opportunistic Networks, *IET Networks*, Vol. 3, No. 3, pp. 228-234, September, 2014.

[24] J. J. Marron, M. A. Labrador, A. Menendez-Valle, D. Fernandez-Lanvin, M. Gonzalez-Rodriguez, Multi Sensor System for Pedestrian Tracking and Activity Recognition in Indoor Environments, *International Journal of Ad Hoc and Ubiquitous Computing*, Vol. 23, No.1-2, pp. 3-23, January, 2016.

[25] R. Berri, D. Wolf, F. Osório, Telepresence Robot with Image-based Face Tracking and 3D Perception with Human Gesture Interface Using Kinect Sensor, *2014 Joint Conference on Robotics: SBR-LARS Robotics Symposium and Robocontrol*, Sao Paulo, Brazil, 2014, pp. 205-210.

[26] J. H. Pepper, *True History of the Ghost: And All about Metempsychosis*, Cambridge University Press, 2012.

[27] J. Grinder, R. Bandler, *The Structure of Magic*, Vol. I & II, Science and Behavior Books, 1976.

[28] Y. M. Huang, C. C. Hsu, Y. N. Su, C. J. Liu, Empowering Classroom Observation with an E-Book Reading Behavior Monitoring System Using Sensing Technologies, *Interacting with Computers*, Vol. 26, No. 4, pp. 372-387, April, 2014.

[29] Y. M. Huang, T. H. Liang, A Technique for Tracking the Reading Rate to Identify the e-book Reading Behaviors and Comprehension Outcomes of Elementary School Students, *British Journal of Educational Technology*, Vol. 46, No. 4, pp. 864-876, June, 2015.

[30] J. Brooke, SUS-A Quick and Dirty Usability Scale, *Usability Evaluation in Industry*, Vol. 189, No. 194, pp. 4-7, June, 1996.

[31] J. P. Chin, V. A. Diehl, K. L. Norman, Development of an Instrument Measuring User Satisfaction of the Human-computer Interface, *CHI '88 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Washington, DC, 1988, pp. 213-218.

[32] H. M. Wen, *The Development of an Affective Tutoring System for Japanese Language Learners*, Master's Thesis, National University of Tainan, Tainan, Taiwan, 2014.

[33] P. Ekman, W. V. Friesen, Constants across Cultures in the Face and Emotion, *Journal of Personality and Social Psychology*, Vol. 17, No. 2, pp. 124-129, February, 1971.

[34] J. A. Russell, A Circumplex Model of Affect, *Journal of Personality and Social Psychology*, Vol. 39, pp. 1161-1178, December, 1980.

[35] R. Bakeman, J. M. Gottman, *Observing Interaction: An Introduction to Sequential Analysis*, Cambridge University Press, 1997.

[36] Y. M. Huang, T. H. Liang, C. H. Chiu, Gender Differences in the Reading of E-books: Investigating Children's Attitudes, Reading Behaviors and Outcomes, *Educational Technology & Society*, Vol. 16, No. 4, pp. 97-110, October, 2013.

[37] R. Shadiev, Y. M. Huang, Facilitating Cross-cultural Understanding with Learning Activities Supported by Speech-to-text Recognition and Computer-aided Translation, *Computers & Education*, Vol. 98, pp. 130-141, July, 2016.

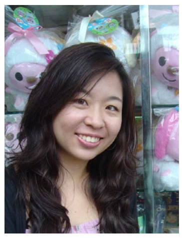
[38] R. Shadiev, T. T. Wu, Y. M. Huang, Enhancing Learning Performance, Attention, and Meditation Using a Speech-to-text Recognition Application: Evidence from Multiple Data Sources, *Interactive Learning Environments*, Vol. 25, No. 2, pp. 249-261, February, 2017.

Biographies

Hao-Chiang Koong Lin works as Dean of College of Office of Student Affairs and also a Professor in the Department of Information and Learning Technology, National University of Tainan. He received his Ph.D. degree in Computer Science from National Tsing-Hua University, 1997. He has published more than 300 internationally refereed research papers focused on affective computing, learning and education technology, digital arts, interaction design, and artificial intelligence.

Wen-Chun Hsu is currently a Ph.D. candidate of the Department of Information and Learning Technology in National University of Tainan in Taiwan, and also a Lecturer of the Department of Visual Communication Design in Kun Shan University. His main research interests are multimedia applications, digital game-based learning, and mobile learning.

Tao-Hua Wang is a Ph.D. student of Department of Engineering Science Department, National Cheng Kung University of Taiwan. She received her master degree in the Department of Information and Learning Technology, National University of Tainan.



Yu-Chun Ma is a doctoral student in Department of Information and Learning Technology, National University of Tainan, Taiwan. Her major is Affective Computing. She holds double master degree, graduating from Graduate Institute of Performing Arts, National Taiwan Normal University and Department of Information and Learning Technology, National University of Tainan.

Meng-Chun Tsai works as designer of advertising design company of Tainan. She is a Ph.D. student in the Department of Information and Learning Technology, National University of Tainan. Her research focuses on Augmented Reality, Information and Digital learning.